MATHEMATISCHE ZEITSCHRIFT

67 BAND

MATHEMATISCHE ZEITSCHRIFT

UNTER STÄNDIGER MITWIRKUNG VON

E. KAMKE
TÜBINGEN

K. KNOPP†
TÜBINGEN

R. NEVANLINNA
HELSINKI

E. SCHMIDT
BERLIN

F. K. SCHMIDT
HEIDELBERG

HERAUSGEGEBEN VON

H. WIELANDT
TÜBINGEN

WISSENSCHAFTLICHER BEIRAT

W. BLASCHKE L. FEJÉR A. E. INGHAM H. KNESER
W. MAGNUS O. PERRON G. PICKERT W. SÜSS

67. BAND

BERLIN · GÖTTINGEN · HEIDELBERG
SPRINGER-VERLAG
1957

Unveränderter Nachdruck 1971
Springer-Verlag Berlin Heidelberg New York

Inhalt des 67. Bandes

Seite

Nachruf auf Konrad Knopp

Agnew, R. P., Borel transforms of Tauberian series 51

Albrecht, R., Auswahlverfahren bei linearer Limitierung 320

Bagemihl, F., On the set of values assumed by holomorphic functions near essential singularities 49

Basu, S. K., On comparison of the total strength of some Hausdorff methods . . . 303

Bauer, F. L., und K. Samelson, Polynomkerne und Iterationsverfahren 93

Behrend, F. A., Uniformizability and compactifiability of topological spaces . . . 203

Bieberbach, L., Eine Bemerkung zur konformen Abbildung zweifach zusammenhängender Gebiete 99

Brauer, G., Some Summation Matrices of Hausdorff Type 397

Collingwood, E. F., On sets of maximum indetermination of analytic functions . . 377

Eichler, M., Eine Verallgemeinerung der Abelschen Integrale 267

Endler, O., Differentiation in algebraischen Funktionenkörpern von n Variablen . . 413

Grauert, H., und R. Remmert, Singularitäten komplexer Mannigfaltigkeiten und Riemannsche Gebiete 103

Grunsky, H., Über konforme Abbildungen, die gewisse Gebietsfunktionen in elementare Funktionen transformieren. I 129

Grunsky, H., Über konforme Abbildungen, die gewisse Gebietsfunktionen in elementare Funktionen transformieren. II 223

Heinhold, J., Zur konformen Abbildung schlichter Gebiete 133

Henstock, R., The summation by convergence factors of Laplace-Stieltjes integrals outside their half plane of convergence 10

Hofmann, J. E., Zur Entwicklungsgeschichte der Eulerschen Summenformel . . . 139

Huppert, B., Lineare auflösbare Gruppen 479

Jacobs, K., Fastperiodizitätseigenschaften allgemeiner Halbgruppen in Banach-Räumen 83

Jurkat, W. B., Über die Umkehrung des Abelschen Stetigkeitssatzes mit funktionentheoretischen Methoden 211

Laugwitz, D., Zur Differentialgeometrie der Hyperflächen in Vektorräumen und zur affingeometrischen Deutung der Theorie der Finsler-Räume 63

Lense, J., Bemerkungen zur Verwendung der Elementarteiler bei Paaren von nicht zerfallenden Kegelschnitten 147

Lenz, H.: Über die elliptischen Funktionen von Jacobi 153

Lingenberg, R., Zur Einführung von Koordinaten in einer projektiven Ebene mit Hilfe von Endomorphismen transitiver Translationsgruppen 332

Ludford, G. S. S., and S. H. Schot, On Sonic Limit Lines in the Hodograph Method 229

Moppert, K.-F., Funktionenscharen im L_2 474

Naumann, H., Über Vektorsterne und Parallelprojektionen regulärer Polytope . . . 75

Seite

PERRON, O., Ein neuartiges diophantisches Problem 176

RAJAGOPAL, C. T., On a theorem of Frobenius and Knopp for Abel summability . . 310

SALZMANN, H., Topologische projektive Ebenen 436

SCHÄFKE, F. W., Integrale über Produkte von Sphäroidfunktionen 238

SCHÄFKE, F. W., Bemerkungen über Variationsrechnung, implizite Funktionen und Lagrangesche Multiplikatoren 252

SCHAUFFLER, R., Die Assoziativität im Ganzen, besonders bei Quasigruppen . . . 428

SCHMIDT, J., Einfacher, ordinalzahlfreier Beweis für die Wohlordnung der Mächtigkeiten 299

SRIVASTAVA, K. J., Fractional integration and Meijer transform 404

STEUERWALD, R., Über die Perioden regelmäßiger Kettenbrüche für Quadratwurzeln aus ganzen Zahlen. II 181

THOMA, E., Zur Reduktionstheorie in separablen Hilbert-Räumen 1

TIETZE, H., Über Schachturnier-Tabellen 188

WALTER, W., Ganze Lösungen der Differentialgleichung $\Delta^p u = f(u)$ 32

WALTER, W., Verallgemeinerte Laplace-Operatoren und Potentiale 38

WALTER, W., Über die Euler-Poisson-Darboux-Gleichung 361

WESTON, J. D., On integration with respect to a measure 457

Konrad Knopp †

Am 30. April 1957 ist KONRAD KNOPP im 75. Lebensjahr unerwartet auf einer Reise verstorben; er erlag in Annecy bei Genf einem Schlaganfall. Damit ist das Leben eines Mathematikers zu Ende gegangen, der in seltener Weise die Gaben eines fruchtbaren Forschers und mitreißenden Lehrers mit der Anziehungskraft einer warmherzigen, weltoffenen und vornehmen Persönlichkeit vereinte.

KNOPPs Forscherarbeit galt der Reihenlehre und vor allem der Limitierungstheorie, deren Entwicklung er stark beeinflußt hat. Auf weite Kreise wirkten und wirken seine Lehrbücher: Die dreibändige „Einführung in die höhere Mathematik", die durch Bearbeitung des von Mangoldtschen Werkes entstand, die „Theorie und Anwendung der unendlichen Reihen" und besonders die fünf Bändchen über Funktionentheorie.

Im Jahre 1918 begründeten KNOPP, L. LICHTENSTEIN, E. SCHMIDT und I. SCHUR gemeinsam die Mathematische Zeitschrift. Nach dem Tode LICHTENSTEINs übernahm KNOPP 1933 die Herausgabe. Mit der ihm eigenen gewinnenden Verbindung von Festigkeit, Takt und feinem Humor verstand er es, die Schriftleitung durch alle Schwierigkeiten der folgenden zwanzig Jahre zu führen und das Ansehen der Zeitschrift nicht nur zu wahren, sondern noch zu erhöhen. Auch nachdem er 1952 das Amt des Herausgebers abgegeben hatte, war er stets bereit, in Bedarfsfällen wieder einzuspringen und beteiligte sich bis in die letzten Tage rege an der redaktionellen Arbeit. Sein Tod hinterläßt eine schmerzlich empfundene Lücke.

Die Schriftleitung.

THOMA, E.
Math. Zeitschr. Bd. 67, S. 1—9 (1957)

Zur Reduktionstheorie in separablen Hilbert-Räumen

Von
ELMAR THOMA

Der Begriff der verallgemeinerten direkten Summe wurde von J. VON NEUMANN in seiner Arbeit: On Rings of Operators. Reduction Theory, Ann. of Math. **50** (1949), eingeführt. Für den wichtigsten Satz (Theorem III) des 1. Teiles jener Arbeit wird hier ein Beweis gegeben, der direkter vorgeht als der Beweis bei VON NEUMANN. Wir sprechen den Satz übrigens etwas allgemeiner aus.

Im 1. Abschnitt führen wir den Begriff des direkten Integrals, der dem der verallgemeinerten direkten Summe entspricht, in einer für unsere Zwecke bequemen Weise ein. Im 2. Abschnitt betrachten wir Spektralmaße und im 3. Abschnitt führen wir den Beweis des oben erwähnten Satzes durch. Dabei benutzen wir neben grundlegenden Sätzen der Integrationstheorie und der Theorie der Hilbert-Räume nur den Satz von RADON-NIKODYM und den Satz, daß für n Vektoren $x_1, x_2, \ldots, x_n$ eines Hilbert-Raumes die positiv semidefinite Matrix $\|(x_i, x_j)\|$ eine vollständige Invariante bezüglich der Gruppe aller unitären Transformationen ist und daß es zu jeder positiv semidefiniten Matrix $\|a_{ij}\|$, $i, j = 1, \ldots, n$, n Vektoren $x_1, \ldots, x_n$ in jedem Hilbert-Raum einer Dimension $\geq n$ gibt, für die $(x_i, x_j) = a_{ij}$ für $i, j = 1, \ldots, n$ ist. Im 4. Abschnitt beweisen wir noch den dazugehörigen Eindeutigkeitssatz mit denselben Hilfsmitteln.

Wir müssen uns wie J. VON NEUMANN auf *separable* Hilbert-Räume beschränken. Für den nicht separablen Fall ist unsere Methode ungeeignet. Zum Beweis eines entsprechenden Satzes im nicht-separablen Fall sind stärkere Hilfsmittel nötig. Vielleicht kommen wir an anderer Stelle auf diese Frage zurück.

1. Das direkte Integral

Unter einem *Hilbert-Raum* verstehen wir hier wie üblich einen linearen Raum über dem Körper der komplexen Zahlen, in dem ein Skalarprodukt erklärt ist und der bezüglich der dadurch eingeführten Norm vollständig ist. Da wir uns im weiteren Verlauf doch auf separable Hilbert-Räume beschränken müssen, wollen wir gleich von Anfang an nur diesen Fall betrachten. Wenn wir im folgenden von einen Hilbert-Raum sprechen, meinen wir daher stets einen *separablen* Hilbert-Raum, d.h. einen Raum endlicher oder abzählbar unendlicher Dimension. Um Ausnahmen zu vermeiden schließen wir noch den Raum der Dimension Null, d.h. den Raum, der alleine aus dem Nullvektor besteht, aus.

Es sei X eine nicht leere Menge und jedem $x \in X$ sei ein Hilbert-Raum $\mathfrak{H}_x$ zugeordnet. Ferner sei $\mathfrak{R}$ ein σ-Körper von Teilmengen von X mit $\mathfrak{R} \ni X$

und μ ein σ-endliches Maß über $\mathfrak{R}$. (Es gibt also Mengen $X_\nu \in \mathfrak{R}$, $\nu = 1, 2, \ldots$ mit $\mu(X_\nu) < +\infty$ und $\bigcup_{\nu=1}^{\infty} X_\nu = X$.) Der Bequemlichkeit halber nehmen wir noch an, daß $\mu|\mathfrak{R}$ vollständig ist, d.h. jede Teilmenge einer μ-Nullmenge zu $\mathfrak{R}$ gehört, und daß $\mu(X) \neq 0$ ist.

Eine „vektorwertige" Funktion f mit dem Definitionsbereich X nennen wir ein *Vektorfeld*[1]), wenn $f(x) \in \mathfrak{H}_x$ für alle $x \in X$. Eine Menge $\mathfrak{M}$ von Vektorfeldern nennen wir *$\mathfrak{R}$-ausgezeichnet*, wenn gilt: Ein Vektorfeld f gehört genau dann zu $\mathfrak{M}$, wenn $(f(x), g(x))$ $\mathfrak{R}$-meßbar für alle $g \in \mathfrak{M}$ ist[2]).

In einer $\mathfrak{R}$-ausgezeichneten Menge $\mathfrak{M}$ existiert $\int |f(x)|^2 \, d\mu$ stets[3]). Wir betrachten nun die Menge $\mathfrak{S}$ aller Vektorfelder aus $\mathfrak{M}$ mit $\int |f(x)|^2 \, d\mu < +\infty$. In $\mathfrak{S}$ kann man dann durch die Definition

$$(f, g) = \int (f(x), g(x)) \, d\mu$$

ein Skalarprodukt einführen. Man zeigt mit derselben Methode wie J. VON NEUMANN (l. c. S. 413, Beweis von Lemma 2), daß $\mathfrak{S}$ ein Hilbert-Raum ist. Dabei muß man natürlich in $\mathfrak{S}$ alle Vektorfelder gleichsetzen, die sich nur auf einer μ-Nullmenge voneinander unterscheiden. Wir nennen diesen Hilbert-Raum $\mathfrak{S}$ das *direkte Integral* der $\mathfrak{H}_x$ mit dem Maße $\mu|\mathfrak{R}$ und der ausgezeichneten Menge $\mathfrak{M}$ und schreiben dafür kurz $\mathfrak{S} = (\int \mathfrak{H}_x \, d\mu; \mathfrak{M})$.

Wenn man die $\mathfrak{H}_x$ auf einer μ-Nullmenge abändert, entsteht in naheliegender Weise eine neue ausgezeichnete Menge $\mathfrak{M}'$ und dazu ein direktes Integral $\mathfrak{S}'$. Ordnet man jedem Vektorfeld $f \in \mathfrak{S}$ ein Vektorfeld $f' \in \mathfrak{S}'$ zu, das sich nur auf einer μ-Nullmenge von f unterscheidet, dann erhält man offensichtlich eine unitäre Abbildung von $\mathfrak{S}$ auf $\mathfrak{S}'$. Wir werden daher im folgenden direkte Integrale als gleich betrachten, die durch Änderung der $\mathfrak{H}_x$ auf einer μ-Nullmenge auseinander hervorgehen.

2. Das zugehörige Spektralmaß

Ist $\mathfrak{H}$ ein Hilbert-Raum und $\mathfrak{R}$ ein σ-Körper von Teilmengen einer nicht leeren Menge X mit $\mathfrak{R} \ni X$ und ist jedem $M \in \mathfrak{R}$ eine Projektion[4]) $E(M)$ in $\mathfrak{H}$ zugeordnet, so nennen wir diese Funktion ein *Spektralmaß* in $\mathfrak{H}$, wenn sie die folgenden Bedingungen erfüllt:

1. $E(X) = I$.

2. Aus $M_\nu \in \mathfrak{R}$, $\nu = 1, 2, \ldots$ und $M_\nu \cap M_\mu = 0$ für $\nu \neq \mu$ folgt $E\left(\bigcup_{\nu=1}^{\infty} M_\nu\right) f = \sum_{\nu=1}^{\infty} E(M_\nu) f$ für alle $f \in \mathfrak{H}$.

[1]) Zu dieser Bezeichnung vgl. R. GODEMENT, Sur la theorie des representations unitaires, Ann. of Math. **53**, S. 79 (1951).

[2]) Mit $(f(x), g(x))$ bezeichnen wir das Skalarprodukt in $\mathfrak{H}_x$.

[3]) $|f(x)| = +\sqrt{(f(x), f(x))}$.

[4]) Eine Projektion E ist eine beschränkte lineare Transformation in $\mathfrak{H}$ mit $E = E^2$ und $E^* = E$; z. B. sind die identische Transformation I $(f \to f)$ und die Nulltransformation 0 $(f \to 0)$ Projektionen.

Durch Hinzunahme aller E-Nullmengen kann man ein Spektralmaß genau so vervollständigen wie ein gewöhnliches Maß. Wir können daher im folgenden stets annehmen, daß ein Spektralmaß vollständig ist.

Es sei jetzt $\mathfrak{S} = (\int \mathfrak{H}_x \, d\mu; \mathfrak{M})$. Jedem $M \in \mathfrak{R}$ ordnen wir die Projektion $E(M)$ mit $E(M) f = \chi_M f$ zu; dabei ist χ_M die charakteristische Funktion von M, d.h. $\chi_M(x) = 0$ für $x \notin M$ und $\chi_M(x) = 1$ für $x \in M$. Man verifiziert sofort, daß diese Funktion $E/\mathfrak{R}$ ein Spektralmaß in $\mathfrak{S}$ ist. Wir nennen $E/\mathfrak{R}$, das zu $\mathfrak{S} = (\int \mathfrak{H}_x \, d\mu; \mathfrak{M})$ *gehörige Spektralmaß*. Es gilt dann: *$E/\mathfrak{R}$ und $\mu/\mathfrak{R}$ sind äquivalent*, d.h. sie haben dieselben Nullmengen. Insbesondere sind daher *$E/\mathfrak{R}$ und $\mu/\mathfrak{R}$ beide gleichzeitig vollständig*.

Beweis. a) Aus $M \in \mathfrak{R}$ und $\mu(M) = 0$ folgt $|E(M) f|^2 = |\chi_M f|^2 = \int\limits_M |f(x)|^2 d\mu = 0$ für alle $f \in \mathfrak{S}$, also ist $E(M) = 0$.

b) Es sei jetzt $M \in \mathfrak{R}$ und $E(M) = 0$. Wir nehmen an, daß $\mu(M) \neq 0$ wäre. Wegen der σ-Endlichkeit von μ gibt es ein $M' \in \mathfrak{R}$ mit $M' \subset M$ und $0 < \mu(M') < +\infty$. Wir wählen ein Vektorfeld g mit $|g(x)| = \chi_{M'}(x)$. Dann ist $g \in \mathfrak{M}$. Wir haben zu zeigen, daß $(f(x), g(x))$ $\mathfrak{R}$-meßbar für alle $f \in \mathfrak{M}$ ist. Da $|g(x)| = \chi_{M'}(x)$ ist, können wir uns auf $f \in \mathfrak{M}$ mit $f(x) = 0$ für $x \notin M'$ beschränken. Wir setzen nun

$$f_n(x) = \begin{cases} f(x) & \text{für } x \text{ mit } |f(x)| \leq n \\ 0 & \text{für } x \text{ mit } |f(x)| > n. \end{cases}$$

f_n gehört zu $\mathfrak{S}$, da $\mu(M') < +\infty$. Wegen $(f_n(x), g(x)) = (E(M) f_n(x), g(x)) = 0$ μ-fast überall, (da $E(M) = 0$), ist $(f_n(x), g(x))$ $\mathfrak{R}$-meßbar und wegen $f(x) = \lim\limits_{n \to \infty} f_n(x)$ ist $(f(x), g(x)) = \lim\limits_{n \to \infty} (f_n(x), g(x))$ und daher $(f(x), g(x))$ $\mathfrak{R}$-meßbar. Da $0 < \int |g(x)|^2 \, d\mu = \mu(M') < +\infty$ ist, gilt $g \in \mathfrak{S}$ und $g \neq 0$. Es ist aber $E(M) g = \chi_M g = g \neq 0$ und daher $E(M) \neq 0$. Das ist ein Widerspruch zur Voraussetzung $E(M) = 0$. Aus $E(M) = 0$ folgt daher $\mu(M) = 0$, w. z. b. w.

Wir sagen ein Hilbert-Raum $\mathfrak{H}$ wird durch das direkte Integral $\mathfrak{S} = (\int \mathfrak{H}_x \, d\mu; \mathfrak{M})$ *dargestellt*, wenn eine unitäre Abbildung U von $\mathfrak{S}$ auf $\mathfrak{H}$ gegeben ist. Das zu $\mathfrak{S}$ gehörige Spektralmaß $E'/\mathfrak{R}$ geht durch $UE'(M) U^{-1} = E(M)$ in ein Spektralmaß $E/\mathfrak{R}$ in $\mathfrak{H}$ über. Wir nennen $E/\mathfrak{R}$ das *zur Darstellung* $(\int \mathfrak{H}_x \, d\mu; \mathfrak{M}; U)$ *gehörige Spektralmaß*. Es ist auch $E/\mathfrak{R}$ zu $\mu/\mathfrak{R}$ äquivalent, da $\mu/\mathfrak{R}$ und $E'/\mathfrak{R}$ und ebenso $E'/\mathfrak{R}$ und $E/\mathfrak{R}$ äquivalent sind.

3. Der Existenzsatz

Es sei ein Hilbert-Raum $\mathfrak{H}$ und ein (vollständiges) Spektralmaß $E/\mathfrak{R}$ in $\mathfrak{H}$ gegeben. Gibt es dann eine Darstellung $(\int \mathfrak{H}_x \, d\mu; \mathfrak{M}; U)$ von $\mathfrak{H}$ mit $E/\mathfrak{R}$ als zugehöriges Spektralmaß? Für separable Hilbert-Räume beantwortet diese Frage der folgende

Satz. *Ist $\mathfrak{H}$ ein separabler Hilbert-Raum und $E/\mathfrak{R}$ ein Spektralmaß in $\mathfrak{H}$, so gibt es*

1. ein endliches Maß $\mu/\mathfrak{R}$, das zu $E/\mathfrak{R}$ äquivalent ist und

2. zu jedem σ-endlichen Maß $\mu/\mathfrak{R}$, das zu $E/\mathfrak{R}$ äquivalent ist, eine Darstellung $(\int \mathfrak{H}_x \, d\mu; \mathfrak{M}; U)$ von $\mathfrak{H}$, zu der das gegebene Spektralmaß $E/\mathfrak{R}$ gehört.

Beweis. Zu 1. Es sei $\Phi_1, \Phi_2, \ldots$ eine orthonormierte Basis von $\mathfrak{H}$ [5]). Ferner seien α_i, $i=1, 2, \ldots$, reelle Zahlen mit $\alpha_i > 0$ und $\sum_{i=1}^{\infty} \alpha_i < +\infty$. Für $M \in \mathfrak{R}$ setzen wir

$$\mu(M) = \sum_{i=1}^{\infty} |E(M)\, \Phi_i|^2 \alpha_i,$$

dann ist $\mu/\mathfrak{R}$ ein Maß, das die 1. Behauptung erfüllt.

Zu 2. Wir führen den Beweis in mehreren Schritten.

1. $\Phi_1, \Phi_2, \ldots$ sei wieder eine orthonormierte Basis von $\mathfrak{H}$. Wir betrachten die Funktionen $\mu_{ij}/\mathfrak{R}$, $i, j = 1, 2, \ldots$, die durch $\mu_{ij}(M) = (E(M)\, \Phi_i, \Phi_j)$ definiert sind. Die μ_{ij} sind komplexwertige, beschränkte Funktionen mit dem Definitionsbereich $\mathfrak{R}$ und, da $E/\mathfrak{R}$ ein Spektralmaß ist, sind sie sogar totaladditiv. Aus $\mu(M) = 0$ folgt wegen der Äquivalenz von $\mu/\mathfrak{R}$ und $E/\mathfrak{R}$ auch $\mu_{ij}(M) = 0$. Nach dem Satz von RADON-NIKODYM gibt es daher $\mathfrak{R}$-meßbare Funktionen F_{ij} mit

$$(1) \qquad (E(M)\, \Phi_i, \Phi_j) = \int_M F_{ij}(x)\, d\mu \quad \text{für alle } M \in \mathfrak{R} \text{ und } i, j = 1, 2, \ldots.$$

2. Die Matrizen

$$(2) \qquad \begin{cases} M_n(x) = \|F_{ij}(x)\|_{i,j=1,\ldots,n} & \text{sind bis auf eine } \mu\text{-Nullmenge} \\ N \subset X \text{ alle } \textit{gleichzeitig} \text{ positiv semidefinit.} \end{cases}$$

Beweis. S_n sei die Menge aller $x \in X$, für die $M_n(x)$ nicht positiv semidefinit ist. Wir haben $\mu\left(\bigcup_{n=1}^{\infty} S_n\right) = 0$ zu zeigen. Dazu genügt es $\mu(S_n) = 0$ zu zeigen. Aus $x \in S_n$ folgt: Es gibt ein n-Tupel $(\lambda_1, \ldots, \lambda_n)$ komplexer Zahlen, so daß $\sum_{\nu,\mu=1}^{n} \lambda_\nu \bar{\lambda}_\mu F_{\nu\mu}(x) \not\geqq 0$ ist. Wegen der Stetigkeit in den λ_ν können wir sogar annehmen, daß die λ_ν rational sind, d.h. $\lambda_\nu = \alpha_\nu + i\beta_\nu$ mit α_ν und β_ν rationale Zahlen. Ist $S_{\lambda_1,\ldots,\lambda_n}$ die Menge aller $x \in X$ mit $\sum_{\nu,\mu=1}^{n} \lambda_\nu \bar{\lambda}_\mu F_{\nu\mu}(x) \not\geqq 0$, dann ist also S_n gleich der Vereinigung aller $S_{\lambda_1,\ldots,\lambda_n}$ mit rationalen λ_ν. Da es nur abzählbar viele n-Tupel mit rationalen λ_ν gibt, ist $\mu(S_n) = 0$, wenn $\mu(S_{\lambda_1,\ldots,\lambda_n}) = 0$ für alle rationalen $(\lambda_1, \ldots, \lambda_n)$ ist. Aus (1) folgt

$$0 \leq \left| E(M) \left(\sum_{\nu=1}^{n} \lambda_\nu \Phi_\nu \right) \right|^2 = \left(E(M) \left(\sum_{\nu=1}^{n} \lambda_\nu \Phi_\nu \right), \sum_{\nu=1}^{n} \lambda_\nu \Phi_\nu \right)$$

$$= \sum_{\nu,\mu=1}^{n} \lambda_\nu \bar{\lambda}_\mu (E(M)\, \Phi_\nu, \Phi_\mu) = \int_M \sum_{\nu,\mu=1}^{n} \lambda_\nu \bar{\lambda}_\mu F_{\nu\mu}(x)\, d\mu$$

für alle $M \in \mathfrak{R}$. Daher kann $S_{\lambda_1,\ldots,\lambda_n}$ nur eine μ-Nullmenge sein. Setzen wir $F_{ij}(x) = 0$ für alle $x \in \bigcup_{n=1}^{\infty} S_n$, dann bleibt (1) richtig und (2) gilt ohne Ausnahme. Wir können daher im folgenden annehmen, daß (2) ohne Ausnahme gilt.

[5]) Hat $\mathfrak{H}$ eine endliche Dimension, so ist $\Phi_1, \Phi_2, \ldots$ als endliche Serie zu betrachten. Im folgenden werden wir uns stets auf den abzählbar-unendlichen Fall beschränken. Die für den endlichen Fall notwendigen Änderungen sind offensichtlich.

3. Mit N bezeichnen wir vorübergehend die Menge aller $x \in X$ mit $F_{i,j}(x) = 0$ für alle $i, j = 1, 2, \ldots$. N gehört zu $\mathfrak{R}$, da alle $F_{i,j}$ $\mathfrak{R}$-meßbar sind. Aus (1) folgt $|E(N)\Phi_i|^2 = \int_N F_{i,i}(x)\,d\mu = 0$. Also ist $E(N) = 0$ und daher auch $\mu(N) = 0$. Auf N können wir also die $F_{i,j}$ nochmals so abändern, daß (1) und (2) ohne Ausnahme erfüllt bleiben und

(3) zu jedem $x \in X$ ein $F_{i,j}$ existiert mit $F_{i,j}(x) \neq 0$.

Man braucht dazu z.B. für $x \in N$ nur $F_{11}(x) = 1$ und $F_{i,j}(x) = 0$ für $i \neq 1, j \neq 1$ setzen. [Im übrigen soll (3) nur bewirken, daß die später zu konstruierenden Hilbert-Räume $\mathfrak{H}_x$ nicht nur aus dem Nullvektor bestehen. Nach der Bemerkung am Schluß des 1. Abschnittes kommt es auf die spezielle Wahl der $F_{i,j}(x)$ auf einer μ-Nullmenge gar nicht an.] Wir nehmen im folgenden an, daß (1), (2) und (3) ohne Ausnahme erfüllt seien.

4. Wir ordnen nun jedem $x \in X$ einen Hilbert-Raum $\mathfrak{H}'_x$ von abzählbarer Dimension zu und konstruieren mittels vollständiger Induktion eine Folge $\varphi_1(x), \varphi_2(x), \ldots$ von Vektoren aus $\mathfrak{H}'_x$ mit

(4) $$(\varphi_i(x), \varphi_j(x)) = F_{i,j}(x) \quad \text{für } i, j = 1, 2, \ldots.$$

$\varphi_1(x), \ldots, \varphi_n(x)$ seien bereits gefunden, so daß (4) für $i, j = 1, \ldots, n$ erfüllt ist. Da $M_{n+1}(x)$ nach (2) positiv semidefinit ist, gibt es $n+1$ Vektoren $\varphi'_1(x), \ldots, \varphi'_{n+1}(x)$ mit $(\varphi'_i(x), \varphi'_j(x)) = F_{i,j}(x)$. Also ist $(\varphi_i(x), \varphi_j(x)) = (\varphi'_i(x), \varphi'_j(x))$ für $i = 1, \ldots, n$ und daher gibt es eine unitäre Transformation V in $\mathfrak{H}'_x$ mit $V\varphi'_i(x) = \varphi_i(x)$, $i = 1, \ldots, n$. Wir setzen $\varphi_{n+1}(x) = V\varphi'_{n+1}(x)$, dann erfüllt $\varphi_1(x), \ldots, \varphi_{n+1}(x)$ (4) für $i = 1, \ldots, n+1$.

5. Wir ordnen nun jedem $x \in X$ den von $\varphi_1(x), \varphi_2(x), \ldots$ in $\mathfrak{H}'_x$ aufgespannten Unterraum $\mathfrak{H}_x$ zu. Wegen (3) und (4) sind nicht alle $\varphi_i(x) = 0$, d.h. $\mathfrak{H}_x \neq 0$ für alle $x \in X$.

$\mathfrak{M}$ sei nun die Menge aller Vektorfelder f mit $f(x) \in \mathfrak{H}_x$ und $(f(x), \varphi_i(x))$ $\mathfrak{R}$-meßbar für $i = 1, 2, \ldots$. Unser nächstes Ziel ist es zu zeigen, daß $\mathfrak{M}$ *eine $\mathfrak{R}$-ausgezeichnete Menge ist.* Wegen (4) ist $\varphi_i \in \mathfrak{M}$. Also folgt aus $(f(x), g(x))$ $\mathfrak{R}$-meßbar für alle $g \in \mathfrak{M}$ insbesondere $(f(x), \varphi_i(x))$ $\mathfrak{R}$-meßbar für $i = 1, 2, \ldots$ und damit $f \in \mathfrak{M}$. Wir haben nun noch zu zeigen: Aus $f \in \mathfrak{M}$ und $g \in \mathfrak{M}$ folgt $(f(x), g(x))$ ist $\mathfrak{R}$-meßbar. Dazu zeigen wir zunächst: Ist $f \in \mathfrak{M}$ und $P_n(x)$ die Projektion auf den von $\varphi_1(x), \ldots, \varphi_n(x)$ aufgespannten Teilraum, dann gibt es $\mathfrak{R}$-meßbare Funktionen α_i mit

(5) $$f_n(x) = P_n(x) f(x) = \sum_{i=1}^{n} \alpha_i(x)\, \varphi_i(x) \quad \text{für alle } x \in X.$$

Beweis. $\{i_1, \ldots, i_r\}$ sei eine Teilmenge von $\{1, \ldots, n\}$, dann ist die Menge $S_{i_1, \ldots, i_r}$ aller $x \in X$, für die die $\varphi_{i_1}(x), \ldots, \varphi_{i_r}(x)$ linear unabhängig sind und denselben Teilraum wie $\varphi_1(x), \ldots, \varphi_n(x)$ aufspannen, wegen (4) $\mathfrak{R}$-meßbar. Es genügt daher (5) für $\mathfrak{R}$-meßbare Teilmengen der Mengen $S_{i_1, \ldots, i_r}$ zu beweisen. Man setze $\alpha_i \equiv 0$ für $i \notin \{i_1, \ldots, i_r\}$ und berechne die übrigen α_i aus

dem Gleichungssystem

$$(f(x), \varphi_{i_\nu}(x)) = \sum_{\mu=1}^{r} \alpha_{i_\mu}(x)\,(\varphi_{i_\mu}(x), \varphi_{i_\nu}(x)), \quad \nu = 1, \ldots, r.$$

Wegen $x \in S_{i_1, \ldots, i_r}$ ist die Determinante dieses Systems $\neq 0$ und wegen $f \in \mathfrak{M}$ und (4) sind alle Koeffizienten $\mathfrak{R}$-meßbar. Also sind auch die Lösungen $\alpha_{i_1}, \ldots, \alpha_{i_r}$ $\mathfrak{R}$-meßbar. (5) ist damit bewiesen.

Es sei nun $f \in \mathfrak{M}$ und $g \in \mathfrak{M}$ und $g_n(x) = P_n(x)\, g(x)$, dann gilt $(f(x), g(x)) = \lim_{n\to\infty} (f_n(x), g_n(x))$ wegen $\lim_{n\to\infty} |f(x) - f_n(x)| = 0$ und $\lim_{n\to\infty} |g(x) - g_n(x)| = 0$. Wegen (5) und (4) ist $(f_n(x), g_n(x))$ $\mathfrak{R}$-meßbar und damit ist dann auch $(f(x), g(x))$ $\mathfrak{R}$-meßbar.

6. Da $\mathfrak{M}$ eine $\mathfrak{R}$-ausgezeichnete Menge ist, können wir das direkte Integral $\mathfrak{E} = (\int \mathfrak{H}_x d\mu; \mathfrak{M})$ bilden. Wegen (1) und (4) ist $(\varphi_i, \varphi_j) = \int_X (\varphi_i(x), \varphi_j(x))\, d\mu$ $= (\boldsymbol{\Phi}_i, \boldsymbol{\Phi}_j) = \delta_{ij}$. Also ist $\varphi_1, \varphi_2, \ldots$ ein orthonormiertes System in $\mathfrak{E}$. Wir zeigen nun:

(6) $\varphi_1, \varphi_2, \ldots$ ist eine Basis in $\mathfrak{E}$.

Beweis. $\tilde{\mathfrak{E}}$ sei der von $\varphi_1, \varphi_2, \ldots$ aufgespannte Unterraum von $\mathfrak{E}$. Wir haben $\tilde{\mathfrak{E}} = \mathfrak{E}$ zu zeigen.

a) $\chi_M \varphi_i \in \tilde{\mathfrak{E}}$ für alle $M \in \mathfrak{R}$, denn es gilt $|\chi_M \varphi_i|^2 = \int_M (\varphi_i(x), \varphi_i(x))\, d\mu =$ $|E(M)\, \boldsymbol{\Phi}_i|^2$ [nach (1) und (4)] $= \sum_{j=1}^{\infty} |(E(M)\, \boldsymbol{\Phi}_i, \boldsymbol{\Phi}_j)|^2$ (da $\boldsymbol{\Phi}_1, \boldsymbol{\Phi}_2, \ldots$ Basis von $\mathfrak{H}$) $= \sum_{j=1}^{\infty} \left| \int_M (\varphi_i(x), \varphi_j(x))\, d\mu \right|^2$ [nach (1) und (4)] $= \sum_{j=1}^{\infty} |(\chi_M \varphi_i, \varphi_j)|^2$. Also ist $\chi_M \varphi_i = \sum_{j=1}^{\infty} (\chi_M \varphi_i, \varphi_j) \varphi_j \in \tilde{\mathfrak{E}}$.

b) Aus $\alpha \varphi_i \in \mathfrak{E}$ und α $\mathfrak{R}$-meßbar folgt $\alpha \varphi_i \in \tilde{\mathfrak{E}}$. α ist nämlich Limes einer Folge $\{\alpha_n\}$ von $\mathfrak{R}$-meßbaren Treppenfunktionen (d.h. Funktionen, die nur endlich viele Werte annehmen) mit $|\alpha_n(x)| \leq |\alpha(x)|$. Nach a) ist $\alpha_n \varphi_i \in \tilde{\mathfrak{E}}$ und wegen $\lim_{n\to\infty} |\alpha \varphi_i - \alpha_n \varphi_i|^2 = \lim_{n\to\infty} \int |\alpha(x) - \alpha_n(x)|^2\, |\varphi_i(x)|^2\, d\mu = 0$ ist auch $\alpha \varphi_i \in \tilde{\mathfrak{E}}$.

c) Es sei nun $f \in \mathfrak{E}$, dann ist wegen (5) auch $f_n \in \mathfrak{M}$ und wegen $|f_n(x)| \leq |f(x)|$ sogar $f_n \in \mathfrak{E}$. Aus (5) und b) folgt $f_n \in \tilde{\mathfrak{E}}$ und wegen

$$\lim |f - f_n|^2 = \lim_{n\to\infty} \int |f(x) - f_n(x)|^2\, d\mu = \int \lim_{n\to\infty} |f(x) - f_n(x)|^2\, d\mu = 0$$

ist dann auch $f \in \tilde{\mathfrak{E}}$. Also ist $\tilde{\mathfrak{E}} = \mathfrak{E}$, w. z. b. w.

7. Wegen (6) gibt es genau eine unitäre Abbildung U von $\mathfrak{E}$ auf $\mathfrak{H}$ mit

(7) $U \varphi_i = \boldsymbol{\Phi}_i$ für $i = 1, 2, \ldots$.

Sei nun $E'/\mathfrak{R}$ das zu $\mathfrak{S} = (\int \mathfrak{H}_x \, d\mu; \mathfrak{M})$ gehörige Spektralmaß, so gilt

$$U E'(M) U^{-1} \Phi_i = U E'(M) \varphi_i = U \chi_M \varphi_i = U \Big(\sum_{j=1}^{\infty} (\chi_M \varphi_i, \varphi_j) \varphi_j \Big)$$

$$= \sum_{j=1}^{\infty} (\chi_M \varphi_i, \varphi_j) U \varphi_j = \sum_{j=1}^{\infty} (\chi_M \varphi_i, \varphi_j) \Phi_j = \sum_{j=1}^{\infty} \int_M (\varphi_i(x), \varphi_j(x)) \, d\mu \cdot \Phi_j$$

$$= \sum_{j=1}^{\infty} (E(M) \Phi_i, \Phi_j) \Phi_j = E(M) \Phi_i .$$

Also ist $U E'(M) U^{-1} \Phi_i = E(M) \Phi_i$ für $i = 1, 2, \ldots$ und da beide Transformationen in $\mathfrak{H}$ beschränkt und linear sind und $\Phi_1, \Phi_2, \ldots$ eine Basis in $\mathfrak{H}$ bilden, gilt also $U E'(M) U^{-1} = E(M)$ für alle $M \in \mathfrak{R}$. Damit ist gezeigt, daß die Darstellung $(\int \mathfrak{H}_x d\mu; \mathfrak{M}; U)$ von $\mathfrak{H}$ die 2. Behauptung unseres Satzes erfüllt.

4. Der Eindeutigkeitssatz

Es seien $(\int \mathfrak{H}_x^i d\mu_i; \mathfrak{M}_i; U_i)$, $i = 1, 2$, zwei Darstellungen von $\mathfrak{H}$, zu denen das Spektralmaß $E/\mathfrak{R}$ gehört. In welcher Beziehung stehen diese beiden Darstellungen? Darüber gibt der folgende Satz Auskunft.

Satz. *Ist $\mathfrak{H}$ ein separabler Hilbert-Raum und $E/\mathfrak{R}$ ein Spektralmaß in $\mathfrak{H}$ und sind $(\int \mathfrak{H}_x^i d\mu_i; \mathfrak{M}_i; U_i)$, $i = 1, 2$, zwei Darstellungen von $\mathfrak{H}$, zu denen das Spektralmaß $E/\mathfrak{R}$ gehört, so gilt:*

1. *$\mu_i/\mathfrak{R}$ ist äquivalent zu $E/\mathfrak{R}$ für $i = 1, 2$. $\mu_1/\mathfrak{R}$ und $\mu_2/\mathfrak{R}$ sind äquivalent und es gibt eine $\mathfrak{R}$-meßbare Funktion γ mit dem Definitionsbereich X mit*

   a) *$\gamma(x) > 0$ für alle $x \in X$ und*

   b) *$\mu_2(M) = \int_M \gamma(x) \, d\mu_1$ für alle $M \in \mathfrak{R}$.*

2. *Bis auf die Punkte einer E-Nullmenge N gibt es zu jedem $x \in X$ eine unitäre Abbildung $V(x)$ von $\mathfrak{H}_x^1$ auf $\mathfrak{H}_x^2$ mit folgenden Eigenschaften:*

   a) *Ist $f_1 \in \mathfrak{S}_1 = (\int \mathfrak{H}_x^1 d\mu_1; \mathfrak{M}_1)$ und*

$$f_2(x) = \begin{cases} \dfrac{1}{\sqrt{\gamma(x)}} \cdot V(x) f_1(x) & \text{für } x \notin N \\ 0 & \text{für } x \in N \end{cases}, \quad \text{so ist } f_2 \in \mathfrak{S}_2 = (\int \mathfrak{H}_x^2 d\mu_2; \mathfrak{M}_2).$$

   b) *Die in* a) *durch $f_1 \to f_2$ definierte Abbildung V ist eine unitäre Abbildung von $\mathfrak{S}_1$ auf $\mathfrak{S}_2$, die die zugehörigen Spektralmaße ineinander überführt.*

Wir bemerken noch: *Aus* 2. *folgt insbesondere, daß $\mathfrak{H}_x^1$ und $\mathfrak{H}_x^2$ bis auf eine E-Nullmenge überall die gleiche Dimension haben.*

Beweis. Zu 1. Am Schluß des 2. Abschnittes haben wir bereits gezeigt, daß $\mu_i/\mathfrak{R}$ und $E/\mathfrak{R}$ äquivalent sind[6]). Die übrigen Behauptungen von 1. sind einfache Folgerungen (Satz von RADON-NIKODYM) aus dieser Tatsache.

Zu 2. 1. $\Phi_1, \Phi_2, \ldots$ sei eine orthonormierte Basis von $\mathfrak{H}$. Ferner sei $\varphi_j^i = U_i^{-1} \Phi_j$, $j = 1, 2, \ldots$. $\varphi_1^i, \varphi_2^i, \ldots$ ist eine orthonormierte Basis von $\mathfrak{S}_i$.

[6]) Im ganzen Beweis ist $i = 1, 2$.

Es sei $\tilde{\mathfrak{H}}_x^i$ der von $\varphi_1^i(x), \varphi_2^i(x), \ldots$ aufgespannte Teilraum von $\mathfrak{H}_x^i$. Wir zeigen nun, daß bis auf eine E-Nullmenge

$$\tilde{\mathfrak{H}}_x^i = \mathfrak{H}_x^i$$

ist. Es sei vorübergehend N_i die Menge aller $x \in X$ mit $\tilde{\mathfrak{H}}_x^i \neq \mathfrak{H}_x^i$. Es genügt zu zeigen, daß N_i eine μ_i-Nullmenge ist. X_j, $j = 1, 2, \ldots$ sei eine Überdeckung von X mit $X_j \in \mathfrak{R}$ und $\mu_i(X_j) < +\infty$. Dann ist N_i sicher eine μ_i-Nullmenge, wenn alle $N_j^i = N_i \cap X_j$ μ_i-Nullmengen sind; das soll nun bewiesen werden.

f_0 sei ein Vektorfeld mit $f_0(x) = 0$ für $x \notin N_j^i$ und $f_0(x) \perp \tilde{\mathfrak{H}}_x^i$ und $|f_0(x)| = 1$ für $x \in N_j^i$. In 2. wird gezeigt, daß $f_0 \in \mathfrak{M}_i$ ist. Wegen $\int |f_0(x)|^2 \, d\mu_i \leq \mu_i(X_j) < +\infty$ ist dann $f_0 \in \mathfrak{S}_i$. Da $(f_0, \varphi_j^i) = \int (f_0(x), \varphi_j^i(x)) \, d\mu_i = 0$ für $j = 1, 2, \ldots$ ist, ist $f_0 = 0$ und damit $\int |f_0(x)|^2 \, d\mu_i = \mu_i(N_j^i) = 0$.

2. Wir haben noch $f_0 \in \mathfrak{M}_i$ zu beweisen. Zunächst ist $(f_0(x), \varphi_j^i(x))$ $\mathfrak{R}$-meßbar $f_0 = 0$ für $j = 1, 2, \ldots$, da $(f_0(x), \varphi_j^i(x)) = 0$ ist. Es sei nun $g \in \mathfrak{S}_i$, dann ist

$$\lim_{n \to \infty} \Big| g - \sum_{j=1}^{n} (g, \varphi_j^i) \varphi_j \Big|^2 = \lim_{n \to \infty} \int \Big| g(x) - \sum_{j=1}^{n} (g, \varphi_j^i) \varphi_j^i(x) \Big|^2 d\mu_i = 0.$$

Man zeigt dann mit den üblichen Methoden, daß es eine Folge $n_1, n_2, \ldots$ gibt mit $g(x) = \lim_{\nu \to \infty} \sum_{j=1}^{n_\nu} (g, \varphi_j^i) \varphi_j^i(x)$ μ_i-fast überall. Wegen $(f_0(x), g(x)) = \lim_{\nu \to \infty} \Big(f_0(x), \sum_{j=1}^{n_\nu} (g, \varphi_j^i) \varphi_j^i(x)\Big)$ μ_i-fast überall ist $(f_0(x), g(x))$ $\mathfrak{R}$-meßbar für alle $g \in \mathfrak{S}_i$.

Es sei jetzt $g \in \mathfrak{M}_i$. Wegen $f_0(x) = 0$ für $x \notin X_j$ können wir uns auf g mit $g(x) = 0$ für $x \notin X_j$ beschränken. Wir setzen

$$g_n(x) = \begin{cases} g(x) & \text{für } x \text{ mit } |g(x)| \leq n \\ 0 & \text{für } x \text{ mit } |g(x)| > n. \end{cases}$$

Wegen $\mu_i(X_j) < +\infty$ gehört dann $g_n \in \mathfrak{S}_i$ und wegen $(f_0(x), g(x)) = \lim_{n \to \infty} (f_0(x), g_n(x))$ ist $(f_0(x), g(x))$ $\mathfrak{R}$-meßbar. Also ist $f_0 \in \mathfrak{M}_i$.

3. $E_i'/\mathfrak{R}$ sei das zu $\mathfrak{S}_i = (\int \mathfrak{H}_x^i \, d\mu_i, \mathfrak{M}_i)$ gehörige Spektralmaß. Da $(\int \mathfrak{H}_x^i \, d\mu_i; \mathfrak{M}_i; U_i)$ eine Darstellung von $\mathfrak{H}$ mit dem dazugehörigen Spektralmaß $E/\mathfrak{R}$ ist, gilt

$$\int_M (\varphi_j^i(x), \varphi_k^i(x)) \, d\mu_i = (E'(M) \varphi_j^i, \varphi_k^i) = (E(M) U_i \varphi_j^i, U_i \varphi_k^i) = (E(M) \Phi_j, \Phi_k).$$

Also ist

$$(8) \quad \begin{cases} \int_M (\varphi_j^1(x), \varphi_k^1(x)) \, d\mu_1 = \int_M (\varphi_j^2(x), \varphi_k^2(x)) \, d\mu_2 \\ = \int_M \gamma(x) (\varphi_j^2(x), \varphi_k^2(x)) \, d\mu_1 \text{ für alle } M \in \mathfrak{R} \text{ und } j, k = 1, 2, \ldots. \end{cases}$$

Aus (8) folgt noch

$$(9) \quad \begin{cases} (\varphi_j^1(x), \varphi_k^1(x)) = \gamma(x) (\varphi_j^2(x), \varphi_k^2(x)) \\ E\text{-fast überall gleichzeitig für } j, k = 1, 2, \ldots. \end{cases}$$

$\tilde{X}$ sei nun die Menge aller $x \in X$ mit

$$(10) \qquad \begin{cases} \tilde{\mathfrak{H}}_x^1 = \mathfrak{H}_x^1, \tilde{\mathfrak{H}}_x^2 = \mathfrak{H}_x^2 \quad \text{und} \quad (\varphi_j^1(x), \varphi_k^1(x)) \\ \quad = \gamma(x)\,(\varphi_j^2(x), \varphi_k^2(x)) \quad \text{für } j, k = 1, 2, \ldots. \end{cases}$$

Nach 1. und 3. ist $N = X - \tilde{X}$ eine E-Nullmenge ($= \mu_1$-Nullmenge $= \mu_2$-Nullmenge). Zu jedem $x \in \tilde{X}$ gibt es wegen (10) genau eine unitäre Abbildung $V(x)$ von $\mathfrak{H}_x^1$ auf $\mathfrak{H}_x^2$ mit

$$(11) \qquad V(x)\,\varphi_j^1(x) = \sqrt{\gamma(x)}\,\varphi_j^2(x), \quad j = 1, 2, \ldots.$$

Jedem Vektorfeld $f_1 \in \mathfrak{S}_1$ ordnen wir nun das (bis auf die μ_2-Nullmenge N bestimmte) Vektorfeld f_2 mit

$$(12) \qquad f_2(x) = \frac{1}{\sqrt{\gamma(x)}} \cdot V(x) f_1(x) \quad \text{für } x \in \tilde{X}$$

zu. Wir schreiben dafür $V f_1 = f_2$ und zeigen, daß V eine unitäre Abbildung des Hilbert-Raumes $\mathfrak{S}_1$ auf den Hilbert-Raum $\mathfrak{S}_2$ ist.

4. Wir werden weiter unten zeigen:

$$(13) \qquad \text{Aus } f_1 \in \mathfrak{S}_1 \text{ folgt } V f_1 \in \mathfrak{S}_2.$$

Es sei $f_1 \in \mathfrak{S}_1$ und $g_1 \in \mathfrak{S}_1$, dann gilt

$$(V f_1, V g_1) = \int \frac{1}{\gamma(x)} \left(V(x) f_1(x),\, V(x) g_1(x)\right) d\mu_2$$

$$= \int \left(f_1(x), g_1(x)\right) \frac{1}{\gamma(x)}\, d\mu_2 = \int \left(f_1(x), g_1(x)\right) d\mu_1 = (f_1, g_1).$$

Also ist V eine unitäre Abbildung von $\mathfrak{S}_1$ in $\mathfrak{S}_2$. Aus (11) folgt aber

$$(14) \qquad V \varphi_j^1 = \varphi_j^2, \quad j = 1, 2, \ldots$$

und da $\{\varphi_j^2\}$ eine Basis von $\mathfrak{S}_2$ ist, ist V eine Abbildung von $\mathfrak{S}_1$ *auf* $\mathfrak{S}_2$.

5. Wir haben noch (13) zu zeigen. Aus (14) [gilt unabhängig von (13)] folgt $V \varphi_j^1 \in \mathfrak{S}_2$ für $j = 1, 2, \ldots$. Entsprechendes gilt auch für alle endlichen Linearkombinationen der φ_j^1. $\mathfrak{S}_1'$ sei die Menge dieser Linearkombinationen. Für $f_1 \in \mathfrak{S}_1'$ und $g_1 \in \mathfrak{S}_1'$ folgt wie oben $(V f_1, V g_1) = (f_1, g_1)$. Ist $f_1 \in \mathfrak{S}_1$, so gibt es eine Folge $\{f_\nu^1\}$ mit $f_\nu^1 \in \mathfrak{S}_1'$ und $\lim_{\nu \to \infty} |f_1 - f_\nu^1| = 0$. Dann gibt es auch eine Teilfolge $\{f_{\nu_k}^1\}$ mit $f_1(x) = \lim_{k \to \infty} f_{\nu_k}^1(x)$ μ_1-fast überall ($= E$-fast überall). Wegen (12) gilt $(V f_1)(x) = \lim_{k \to \infty} (V f_{\nu_k}^1)(x)$ E-fast überall ($= \mu_2$-fast überall). Da aber V zumindest in $\mathfrak{S}_1'$ eine unitäre Abbildung ist, ist $V f_{\nu_k}^1$ in $\mathfrak{S}_2$ eine Cauchy-Folge. Es gibt also ein $f \in \mathfrak{S}_2$ mit $\lim_{k \to \infty} |f - V f_{\nu_k}^1| = 0$ und daher eine Teilfolge $\{f_{\nu_l}^1\}$ von $\{f_{\nu_k}^1\}$ mit $f(x) = \lim_{l \to \infty} (V f_{\nu_l}^1)(x)$ μ_2-fast überall. Also ist $f(x) = (V f_1)(x)$ μ_2-fast überall und daher $f = V f_1 \in \mathfrak{S}_2$.

6. Es ist $V E_1'(M) V^{-1} \varphi_j^2 = V E_1'(M) \varphi_j^1 = V \chi_M \varphi_j^1 = \chi_M \varphi_j^2 = E_2'(M) \varphi_j^2$ für $j = 1, 2, \ldots$. Also gilt $V E_1'(M) V^{-1} = E_2'(M)$ für alle $M \in \mathfrak{R}$. Damit ist unser Satz vollständig bewiesen.

München, Mathematisches Institut der Universität

(Eingegangen am 23. Oktober 1956)

R. HENSTOCK
Math. Zeitschr. Bd. 67, S. 10—34 (1957)

The summation by convergence factors of LAPLACE-STIELTJES integrals outside their half plane of convergence

By
R. HENSTOCK

When $\sum_{n=0}^{\infty} f_n v^n$ is a power series, then the limit as m tends to infinity of the expression $\sum_{n=0}^{\infty} g_{mn} f_n v^n$ some times enables us to sum the power series in regions outside the circle of convergence[1]). Putting $v = e^{-z}$ we see that the power series can be expressed as a LAPLACE-STIELTJES integral

$$\int_0^{\infty} e^{-zw}\, ds(w), \tag{1}$$

where in this special case the only variation of $s(w)$ consists of jumps at the integer values of w and the integral is convergent in a vertical half plane.

For general functions $s(w)$, if (1) is convergent at one point z it is again convergent in a vertical half plane, and the present paper studies the summation of (1) by convergence factors $a(x, w)$ in regions outside the half plane of convergence.

Paper [4] and the formal theory of this paper, except Lemma 1 and Theorem 2, were stimulated by lectures on the corresponding matrix theory given by Dr. R. G. COOKE.

1. Notation

We suppose that $f(w)$ is defined and finite in $w \geq 0$, and we denote by $[f]_{\alpha}^{\beta}$ the difference $f(\beta) - f(\alpha)$, and by $V[f; W]$ the *weak variation*[2]) of $f(w)$ over a set W in $w \geq 0$. If the limits exist,

$$f(w-) = \lim_{u \to w,\ 0 \leq u < w} f(u), \qquad f(w+) = \lim_{u \to w,\ 0 \leq w < u} f(u).$$

To simplify enunciations we assume that $f(w) = f(0+)$ for $w < 0$, whenever $f(0+)$ exists. Then $f(0-) = f(0+)$.

If $f(w)$ is of bounded variation in $w \geq 0$ we define

$$f_c(0) = f(0+), \qquad f_c(w) = f(w-) - \sum_{0 < u < w} \{f(u+) - f(u-)\} \qquad (w > 0),$$

the *continuous part of* $f(w)$.

[1]) See, for example, OKADA [6], COOKE [1], Chapter 8.
[2]) See, for example, SAKS [7], p. 221, § 4.

If W is a set in $w \geq 0$ we denote its LEBESGUE measure by $\mu(W)$ and its characteristic function by $ch(W; w)$. The intervals $\alpha \leq w \leq \beta$ and $\alpha < w < \beta$ are respectively denoted by $[\alpha, \beta]$, (α, β).

The symbols (R), (L), (LS), (PS), (IP), placed before an integral sign denote respectively that the integral is the RIEMANN integral (or a contour integral of RIEMANN type), the LEBESGUE integral, the LEBESGUE-STIELTJES integral, WARD's PERRON-STIELTJES integral, and the "integral by parts" of [4], defined to be

$$(2) \qquad (IP)\int_{\alpha}^{\beta} f(w)\, ds(w) \equiv f(\beta)\, s(\beta) - f(\alpha)\, s(\alpha) - (PS)\int_{\alpha}^{\beta} s(w)\, df(w),$$

whenever the integral on the right exists. This integral (2) is the only integral strong enough for our purpose, as is explained in [4]. When $f(w)$ is of bounded variation and $s(w)$ is a bounded BAIRE[3]) function in $\alpha \leq w \leq \beta$, then[4]) (2) exists and

$$(3) \qquad \begin{cases} (IP)\int_{\alpha}^{\beta} f(w)\, ds(w) = f(\beta-)\, s(\beta) - f(\alpha+)\, s(\alpha) - \\ \qquad - \sum_{\alpha < w < \beta} s(w)\{f(w+) - f(w-)\} - (LS)\int_{\alpha}^{\beta} s(w)\, df_c(w). \end{cases}$$

As in [4], the generalized limit $S(s)$ of a bounded BAIRE function $s(w)$ as w tends to infinity, is taken to be of the form

$$(4) \qquad S(s) \equiv \lim_{x \to \infty} S(s, x) \equiv \lim_{x \to \infty} \lim_{w \to \infty} S(s, x, w),$$

where

$$(5) \qquad S(s, x, w) \equiv (IP)\int_{0}^{w} a(x, u)\, ds(u),$$

and where

$$(6) \qquad \lim_{u \to w} \inf a(x, u) \leq a(x, w) \leq \lim_{u \to w} \sup a(x, u)$$

when w and u are in $w \geq 0$. In particular, if $a(x, 0+)$ exists we assume that

$$a(x, 0-) = a(x, 0) = a(x, 0+).$$

The use of (6) simplifies the integration theory; if it is omitted we need only replace $a(x, w)$ by $a(x, w+)$ in the subsequent theorems, and add conditions corresponding to [5], Theorem 3 (9), (10), (11), where $j(u)$ is as in [5], Theorem 4 (36), with $g(u)$ replaced by $a(x, u)$, and with $a = 0$ and b arbitrarily large.

The LAPLACE-STIELTJES integrals are taken to be of the form

$$(7) \qquad F(s(\cdot); z) \equiv F(z) \equiv \lim_{w \to \infty} F(z, w),$$

[3]) A BAIRE (BOREL-*measurable*) function is any function that can be obtained from continuous functions by using repeated limits.

[4]) See [4], p. 276, (1.6).

where

$$F(z,w) \equiv (IP)\int_0^w e^{-zu}\,ds(u) = e^{-zw}s(w) - s(0) + (L)\int_0^w z e^{-zu} s(u)\,du, \tag{8}$$

and where $s(w)$ is a BAIRE function bounded in each finite interval of $w \geq 0$. The manipulation of (8) needs the following Lemma.

Lemma 1. *If $b(w)$ and $c(w)$ are finite in $w \geq \alpha$, and if $a(w)$ is bounded in each finite interval of $w \geq \alpha$, and if*

$$I(w) \equiv (IP)\int_\alpha^w b(u)\,dc(u), \qquad (IP)\int_\alpha^w a(u)\,dI(u)$$

exist for all $w \geq \alpha$, then

$$(IP)\int_\alpha^w a(u)\,dI(u) = (IP)\int_\alpha^w a(u)\,b(u)\,dc(u).$$

From Lemma 1, taking

$$a(u) = e^{(v-z)u}, \qquad b(u) = e^{-vu}, \qquad c(u) = s(u),$$

we obtain, for all complex z, v,

$$F(z,w) = (IP)\int_0^w e^{(v-z)u}\,d_u F(v,u) \tag{9}$$

$$= e^{(v-z)w}F(v,w) + (L)\int_0^w (z-v)\,e^{(v-z)u}F(v,u)\,du. \tag{10}$$

Let $R(r)$ denote the half plane $re\ z > r$. Then if $F(z)$ is convergent at some point $z = v$, it is also by (10) convergent in $R(re\ v)$, and equal there to

$$F(z) = (L)\int_0^\infty (z-v)\,e^{(v-z)w}F(v,w)\,dw. \tag{11}$$

Thus the integral for $F(z)$, given by (7), (8), is convergent in $R(\sigma)$ for some minimal $\sigma < +\infty$, and divergent at all points of $re\,z < \sigma$. Note that σ can be $-\infty$.

When $s(w)$ has jumps at the integers, and is constant in the intervals between the integers, $R(\sigma)$ corresponds, by the exponential transformation $v = e^{-z}$, to the circle $|v| < e^{-\sigma}$ of convergence of a power series in v. The domain of z that corresponds to the principal star domain of the series, is defined as follows.

If Z is a set of complex numbers we denote its complement by CZ, and for a fixed complex number v we denote by $Z+v$ the set of $z+v$ for all z in Z.

We define an *H-set* to be a set Z of complex numbers with the following properties.

(12) Z is closed.

If v lies in Z then so do all points z with

$$\text{im}\, z = \text{im}\, v, \quad \text{re}\, z \leqq \text{re}\, v. \tag{13}$$

(14) The zero 0 lies in Z.

(15) No point of $R(0)$ lies in Z.

(16) The section of Z by each line $\text{re}\, z = \text{constant}$ is bounded.

Under (13) and (14), Z contains Z_0, the line $\text{im}\, z = 0$, $\text{re}\, z \leqq 0$. For each function $F(z)$ of type (7), (8) we can define a real function $\alpha = \alpha(\beta)$ of the real number β to be such that $F(z)$ can be continued analytically along the whole line $\text{re}\, z \geqq \gamma$, $\text{im}\, z = \beta$, for all $\gamma > \alpha$ but for no $\gamma < \alpha$. Then $-\infty \leqq \alpha \leqq \sigma$ for all β.

If Z is an H-set we now put

$$H(F;Z) \equiv C \bigcup_{\beta} \{Z + (\alpha + i\beta)\}.$$

Then the *principal horizontal domain of* $F(z)$ is $H_0 \equiv H(F;Z_0)$. In words, it consists of lines λ parallel to the real axis, including such part of λ that lies to the right of all singularities on λ, or including the whole of λ if there is no singularity on λ. We suppose that the value of $F(z)$ in H_0 is that given by direct analytic continuation from the half plane of convergence of the integral for $F(z)$.

The *partial horizontal domain*[5]) of $F(z)$, relative to Z, is $H(F;Z)$, and

$$H(F;Z) \subseteqq H_0.$$

For the summation of the LAPLACE-STIELTJES integrals given by (7), (8), we replace $s(u)$ in (5) by $F(z,u)$, and thus we consider

$$B(F;z,x) \equiv (IP) \int_0^\infty a(x,w)\, d_w F(z,w), \tag{17}$$

where $a(x,w)$ is a finite BAIRE function satisfying (6). By Lemma 1, if $a(x,w)$ is bounded in each finite interval of $w \geqq 0$ we obtain

$$B(F;z,x) = (IP) \int_0^\infty a(x,w)\, e^{-zw}\, ds(w). \tag{18}$$

Let Z be an H-set. Taking a fixed real number p, we say that the function $a(x,w)$ is *efficient* $(Z;p)$ for an $F(z)$ satisfying (7), (8), if

(19) $B(F;z,x)$ exists for each $x \geqq 0$ and each z in $R(\sigma - p)$,

and as $x \to \infty$, $B(F;z,x)$ is uniformly convergent to $F(z)$ in each compact set contained in

$$H(F;Z) \cap R(\sigma - p), \tag{20}$$

whenever $F(z)$ is convergent in $R(\sigma)$.

[5]) Compare the definition of partial star domain given in COOKE [1], p. 189.

In the following pages we study the problem of finding necessary and sufficient conditions on $a(x, w)$ in order that it should be *efficient* $(Z; p)$ for all $F(z)$ satisfying (7), (8), or for a special subclass of $F(z)$. Not all functions *efficient* $(Z; p)$ for all $F(z)$ of type (7), (8) satisfy the conditions of [4], p. 276, Theorem 3.A proof of this statement could be constructed following that given in [3]. The reason for this result is that H_0 is an open set. The integral for $F(z)$ may be convergent at a point on the boundary of H, whereas we do not assume that $B(F; z, x)$ tends to $F(z)$ there.

2. A direct definition of the integral by parts, and the proof of Lemma 1.

Given the finite functions $f(w)$, $s(w)$, defined in $[\alpha, \beta]$, we say that $D(w)$ is a *major function of* $f(w)$ *with respect to* $s(w)$ *in* $[\alpha, \beta]$, *in the sense of the integral by parts*, if $D(\alpha) = 0$, and for u, w in $[\alpha, \beta]$, there is a $\delta(w) > 0$ such that

$$[D]_w^u \geq f(u)\,[s]_w^u \quad \text{if} \quad 0 < u - w \leq \delta(w),$$

$$[D]_w^u \leq f(u)\,[s]_w^u \quad \text{if} \quad 0 > u - w \geq -\delta(w).$$

The upper integral by parts of $f(w)$ *with respect to* $s(w)$, is

$$(IP)\int_\alpha^{-\beta} f(w)\,ds(w) \equiv \inf D(\beta)$$

for all such major functions $D(w)$, if there are any. If not, we write $+\infty$ for the value of the upper integral. Now

$$[fs - D]_w^u - s(w)\,[f]_w^u = f(u)\,[s]_w^u - [D]_w^u,$$

so that $D(w)$ is a major function of $f(w)$ with respect to $s(w)$ in $[\alpha, \beta]$, in the sense of the integral by parts, if and only if $[fs]_\alpha^w - D(w)$ is a minor function of $s(w)$ with respect to $f(w)$ in $[\alpha, \beta]$, in the sense of WARD's PERRON-STIELTJES integral[6]). Hence

$$(21) \qquad (IP)\int_\alpha^{-\beta} f(w)\,ds(w) = [fs]_\alpha^\beta - (PS)\int_{-\alpha}^{\beta} s(w)\,df(w).$$

Similarly we can define minor functions in the sense of the integral by parts, and then the lower integral by parts, obtaining

$$(22) \qquad (IP)\int_{-\alpha}^{\beta} f(w)\,ds(w) = [fs]_\alpha^\beta - (PS)\int_\alpha^{-\beta} s(w)\,df(w).$$

If the upper and lower integrals by parts are finite and equal, we define the common value as the *integral by parts*, obtaining from (21) and (22),

$$(23) \qquad (IP)\int_\alpha^{\beta} f(w)\,ds(w) = [fs]_\alpha^\beta - (PS)\int_\alpha^{\beta} s(w)\,df(w),$$

so that the new definition agrees with [4], (1.2), and with the present (2).

To prove Lemma 1 we now follow the proof of [9], p. 588, Theorem 6, taking major and minor functions in the sense of the integral by parts.

[6]) See WARD [9], p. 584, inequalities (5), (6).

3. *The existence of* $B(F; z, x)$

In this section we consider requirement (19) in the definition of *efficiency* $(Z; p)$.

Theorem 1. *If* $a(x, w)$ *is a finite* BAIRE *function satisfying* (6), *then in order that the* $B(F; z, x)$ *of* (17) *should exist for a fixed* $x \geq 0$, *for* z *in* $R(\sigma - p)$, *and for each* $F(z)$ *satisfying* (7), (8), *convergent in* $R(\sigma)$, *and with* $s(w)$ *a* BAIRE *function bounded in each finite interval of* $w \geq 0$, *it is necessary and sufficient that*

$$M_1(q, x) \equiv V[e^{qw} a(x, w); (0, \infty)] < \infty \tag{24}$$

for each $q < p$.

For the necessity we take $s(w)$ to be bounded in $w \geq 0$. Then the integral for $F(z)$ is convergent at least in $R(0)$, so that by hypothesis, $B(F; z, x)$ is convergent at least in $R(-p)$. From (8),

$$F(0, w) = s(w) - s(0),$$

so that we take $z = 0$ and use the proof of [4], p. 275, Theorem 2, first part, to show that $a(x, w)$ is of bounded variation in each finite interval of $w \geq 0$, for the particular $x \geq 0$. Applying Lemma 1 we obtain (18). Hence taking $z = -q$ for $q < p$ we obtain (24) by the proof of [4], p. 276, Theorem 3 (ii), first part.

To prove the sufficiency let z be fixed in $R(\sigma - p)$. Then there is a $q < p$ such that $\operatorname{re} z + q < \sigma$, and thus

$$\lim_{w \to \infty} F(z + q, w) = F(z + q).$$

Taking $w > u \geq 0$, and using (24) and Lemma 1, we obtain

$$|e^{qw} a(\alpha, w) - a(x, 0)| \leq M_1(q, x) < \infty,$$

$$\begin{aligned}
(IP) \int_u^w a(x, y)\, d_y F(z, y) &= (IP) \int_u^w e^{qy} a(x, y)\, d_y \left\{ (IP) \int_u^y e^{-q\eta}\, d_\eta F(z, \eta) \right\} \\
&= (IP) \int_u^w e^{qy} a(x, y)\, d_y \{F(z + q, y) - F(z + q, u)\} \\
&= (IP) \int_u^w e^{qy} a(x, y)\, d_y \{F(z + q, y) - F(z + q)\} \\
&= \left[e^{qy} a(x, y) \{F(z + q, y) - F(z + q)\} \right]_u^w - \\
&\quad - (PS) \int_u^w \{F(z + q, y) - F(z + q)\}\, d_y \{e^{qy} a(x, y)\},
\end{aligned}$$

$$\left| (IP) \int_u^w a(x, y)\, d_y F(z, y) \right| \leq \tag{25}$$

$$\leq \{2 M_1(q, x) + |a(x, 0)|\} \sup_{u \leq y \leq w} |F(z + q, y) - F(z + q)|,$$

and the right-hand side of (25) tends to 0 as $u, w \to \infty$. Hence $B(F; z, x)$ exists as a convergent integral in $R(\sigma - p)$. Further, from (25),

$$(26)\quad |B(F; z, x)| \leq \{2M_1(q, x) + |a(x, 0)|\} \sup_{w \geq 0} |F(z + q, w) - F(z + q)|.$$

This result will be needed later.

4. *A special subclass of the functions* $F(z)$

For real $r \geq 0$, we say that $F(z)$ *belongs to the class* S_r if $F(z)$ satisfies (7), (8), with $s(w)$ a BAIRE function satisfying

$$(27)\qquad s(w) = O(1), \qquad (L)\int_w^{w+1} |s(u)|\, du = O(e^{-rw}),$$

as $w \to \infty$.

Theorem 2. *If $a(x, w)$ is a finite* BAIRE *function satisfying* (6), *and if $p > 0$, then in order that for every $0 \leq r < p$, $a(x, w)$ should be efficient $(Z; r)$ for each $F(z)$ of S_r, it is necessary and sufficient that* (24) *is true for all $x \geq 0$, $q < p$; that*

$$(28)\qquad a(x, w\pm) \text{ tend to } 1 \text{ as } x \to \infty \text{ for each fixed } w \geq 0;$$

$$(29)\qquad \lim_{x \to \infty} \sum_{w > 0} e^{rw} |a(x, w+) - a(x, w-)| = 0,$$

for each $q < p$;

for each $k \geq 0$, and for each BOREL *set W satisfying*

$$(30)\qquad \mu\{W \cap (w, w+1)\} = O(e^{-kw})$$

as $w \to \infty$, we have, for each $q < \inf(p, k)$, the two conditions

$$(31)\qquad \lim_{x \to \infty} (LS)\int_W e^{rw}\, d_w a_c(x, w) = 0,$$

$$(32)\quad (LS)\int_W e^{qw}\, d_w V[a_c(x, u); (0, w)] \leq M_2(q, k)\left[1 + \sup_{w \geq 0} e^{kw}\mu\{W \cap (w, w+1)\}\right],$$

where $M_2(q, k)$ is finite and independent of W and of $x \geq x_0(q, k)$.

It will be shown later that

$$(33)\qquad \limsup_{x \to \infty} M_1(q, x) < \infty \quad \text{for each} \quad q < 0,$$

follows from (28), (29), and (32) for $k = 0$.

Note that for each $w > 0$,

$$\liminf_{x \to \infty} M_1(q, x) \geq \lim_{x \to \infty} \inf |e^{qw} a(x, w-) - a(x, 0+)| = |e^{qw} - 1|.$$

Hence for $0 < q < p$, $M_1(q, x) \to \infty$ as $x \to \infty$, and (33) cannot be true for any $q > 0$.

The necessity of (24) follows from the proof of the necessity in Theorem 1. The necessity of (28) follows as does the necessity of Theorem 3 (i) of [4], p. 276. To prove the necessity of (29), (31), (32) we assume that $s(w)$ satisfies

(27) with $s(0)=0$. Then for $z\neq 0$, integrating by parts in a LEBESGUE integral,

$$\begin{aligned} F_1(z,w) &\equiv (IP)\int_0^w e^{-zu}\,d_u\left\{(L)\int_0^u s(y)\,dy\right\} \\ &= e^{-zw}(L)\int_0^w s(y)\,dy + z(L)\int_0^w\left\{(L)\int_0^u s(y)\,dy\right\}e^{-zu}\,du \\ &= (L)\int_0^w e^{-zu}s(u)\,du \\ &= -z^{-1}e^{-zw}s(w) + z^{-1}F(z,w), \end{aligned}$$

$$F(z,w) = e^{-zw}s(w) + zF_1(z,w). \tag{34}$$

The formula (34) is true also for $z=0$. From (34), for z in $R(0)$, as $s(w)$ is bounded,

$$F(z) = zF_1(z), \quad \text{where} \quad F_1(z) \equiv \lim_{w\to\infty} F_1(z,w). \tag{35}$$

By analytic continuation (35) is true for all z in the principal horizontal domain of $F_1(z)$. In particular, it is true in $R(-r)$, since by (27), the integral defining $F_1(z)$ is absolutely convergent there. Hence from (34), (35), and by hypothesis, $e^{-zw}s(w)$ is summed to 0 in $R(-\inf(p,r))$. Taking $z=-q>-\inf(p,r)$, we see that

$$e^{qw}s(w) \text{ } \textit{is summed to } 0 \textit{ for all } q<\inf(p,r). \tag{36}$$

If (29) is false for some $q<p$, then for some sequence $x_n\to\infty$ and some $\varepsilon>0$,

$$\sum_{w>0} e^{qw}|a(x_n,w+)-a(x_n,w-)| \geq \varepsilon > 0. \tag{37}$$

Let $\{w_m\}$ be the countable set of discontinuities in $w>0$ of the $a(x_n,w)$ for all n. Then if $s_1(w)$ is bounded, and 0 unless w lies in $\{w_m\}$, $s_1(w)$ will satisfy (27) for all r, and by (36), for $r=p$ and $q<p$,

$$\sum_{m=1}^{\infty} v_{nm}t_m \to 0$$

as $n\to\infty$, where

$$v_{nm} \equiv \exp(qw_m)\{a(x_n,w_m+)-a(x_n,w_m-)\}, \qquad t_m \equiv s_1(w_m).$$

Hence by [4], p. 281, Lemma 2,

$$\sum_{m=1}^{\infty}|v_{nm}| \to 0$$

as $n\to\infty$, contradicting (37). Thus (29) is true.

From (24), (27), (29), (36), with $q<\inf(p,r)$ and $s(0)=0$,

$$(IP)\int_0^{\infty} a(x,w)\,d_w\{e^{qw}s(w)\}\to 0,$$

$$\left[e^{qw}s(w)\,a(x,w)\right]_0^{\infty}=0,$$

$$\sum_{w>0} e^{qw}s(w)\{a(x,w+)-a(x,w-)\} = O\Big(\sum_{w>0} e^{qw}|a(x,w+)-a(x,w-)|\Big)\to 0$$

as $x \to \infty$, and therefore

$$(38)\quad \begin{cases} (LS)\int_0^\infty e^{q w}\, s(w)\, d_w a_c(x,w) \\ \qquad = [e^{q w} s(w)\, a(x,w)]_0^\infty - \sum_{w>0} e^{q w} s(w)\{a(x,w+) - a(x,w-)\} - \\ \qquad\quad - (IP)\int_0^\infty a(x,w)\, d_w\{e^{q w} s(w)\} \to 0 \end{cases}$$

as $x \to \infty$, for each $q < \inf(p, r)$. To obtain (31) we have only to take $s(w)$ as the characteristic function of W, with $r = k$.

Let us define the norm $\|s\|$ of a function $s(w)$ satisfying (27) for some fixed r, to be

$$\|s\| = \sup_{w \geq 0} |s(w)| + \sup_{w \geq 0} \left\{ e^{r w}\ (L)\int_w^{w+1} |s(u)|\, du \right\}.$$

The class S_r then becomes a BANACH space[7]), as it is easily shown to be complete.

For each $q < p$, the transformation

$$(39)\qquad (LS)\int_0^\infty e^{q w}\, s(w)\, d_w a_c(x,w)$$

applied to functions $s(w)$ of S_r, is obviously additive and homogeneous[8]). To prove that it is also continuous we need a lemma.

Lemma 2. *If, for each* $q < p$,

$$M_3(q) \equiv V\left[e^{q w} a(w); [0, \infty)\right] < \infty,$$

then, for each $q < p$,

$$|a(w)| \leq e^{-q w} M_3(q).$$

For fixed $q < p$ there is an $\varepsilon > 0$ such that $q + \varepsilon < p$, and then

$$|e^{(q+\varepsilon) w} a(w) - e^{(q+\varepsilon) u} a(u)| \leq V\left[e^{(q+\varepsilon) y} a(y); [u, w]\right] \to 0$$

for $u < w$, as $u \to \infty$. Hence

$$\lim_{w \to \infty} e^{(q+\varepsilon) w} a(w)$$

exists, and so

$$\lim_{w \to \infty} e^{q w} a(w) = 0,$$

$$|e^{q u} a(u)| \leq V\left[e^{q y} a(y); [u, \infty]\right] \leq M_3(q).$$

Hence the lemma.

Now by (24) and Lemma 2, for each $q < p$ and $\varepsilon = \frac{1}{2}(p - q)$,

$$(40)\quad \begin{cases} (LS)\int_0^\infty e^{q w}\, s(w)\, d_w a_c(x,w) \\ \qquad = (LS)\int_0^\infty s(w)\, d_w\{e^{q w} a_c(x,w)\} - (L)\int_0^\infty s(w)\, a_c(x,w)\, q\, e^{q w}\, dw, \\ \left|(LS)\int_0^\infty e^{q w}\, s(w)\, d_w a_c(x,w)\right| \leq \|s\| M_1(q,x) + \|s\|\, q\, M_1(q+\varepsilon, x)/\varepsilon, \end{cases}$$

[7]) See, for example, [10], pp. 92—93.

[8]) [10], pp. 133—134.

so that (39) is continuous. By (38), (40), and the BANACH-STEINHAUS theorem [9],

$$(41) \qquad \left|(LS)\int_0^\infty e^{qw}\, s(w)\, d_w a_c(x,w)\right| \leq M_2(q,r)\,\|s\| < \infty$$

for some $x_0 = x_0(q,r)$ and each x in $x \geq x_0$.

To obtain (32) from (41) we need another lemma.

Lemma 3. *If $a(w)$ is continuous and of bounded variation in $\alpha \leq w \leq \beta$, and if $V(w)$ is its weak variation in (α, w) then there is a finite function $\psi(w)$ defined and measurable with respect to $V(w)$ in $\alpha \leq w \leq \beta$, such that for $\alpha \leq u < w \leq \beta$,*

$$(42) \qquad a(w) - a(u) = (LS)\int_u^w \psi(y)\, dV(y),$$

$$(43) \qquad V[a(w); W] = (LS)\int_W |\psi(y)|\, dV(y),$$

$$(44) \qquad |\psi(w)| = 1 \quad \textit{for all points } w \textit{ of } \alpha \leq w \leq \beta.$$

For let $\{(u_n, w_n)\}$ be a sequence of non-overlapping intervals. Then

$$\sum_{n=1}^{\infty} |a(w_n) - a(u_n)| \leq \sum_{n=1}^{\infty} V[a(w); (u_n, w_n)],$$

so that $a(w)$ is absolutely continuous with respect to $V(w)$. As $a(w) - a(\alpha)$ is continuous and of bounded variation, it is a signed measure of the interval (α, w), in the sense of [2], p. 118. From $a(w)$ can be constructed a signed measure applying to each BOREL set in $[\alpha, \beta]$ and to other sets. The corresponding measure constructed from $V(w)$ is $V[a(w); W]$. It follows by [2], pp. 128–129, § 31, Theorem B, that there is a function $\psi(w)$ satisfying (42) for each u, w in $\alpha \leq u < w \leq \beta$.

From (42) we obtain (43) by [2], p. 123, result (5). As (43) is also true for the function $|\psi(y)| \equiv 1$, we must therefore have $|\psi(y)| = 1$ except in a set Y of y with $V[a(w); Y] = 0$, so that we can suppose that (44) ist rue. This change in ψ will not alter (42) nor (43).

Continuing with the proof of Theorem 2, necessity, we take the special function

$$s_2(w) \equiv \overline{\operatorname{sgn}}\,\{\psi(x,w)\}\, ch(W; w),$$

where

$$\overline{\operatorname{sgn}}\, z \equiv |z|/z \;(z \neq 0), \qquad \overline{\operatorname{sgn}}\, 0 = 0,$$

where $\psi(x,w)$ is the function $\psi(w)$ of Lemma 3 when $a(w)$ is $a_c(x,w)$, and where $ch(W; w)$ is the characteristic function of the set W, satisfying (30). By (44), for s_2 in S_r,

$$\|s_2\| = 1 + \sup_{w \geq 0} e^{-rw}\,\mu\{W \cap (w, w+1)\},$$

so that (41) and (43) give (32).

[9] [10], p. 135, § 2, Theorem 1.

To prove (33), from (24), (28), (29), (32), we first take $k=0$, $W=(0,\infty)$ in (32), obtaining, for $q<0$, $x\geq x_0(q,0)$,

$$(LS)\int_0^\infty e^{qw}\,d_w V(x,w)\leq 2M_2(q,0),$$

$$e^{qw}V(x,w)\leq 2M_2(q,0),$$

$$|a_c(x,w)|\leq |a_c(x,0)|+2M_2(q,0)\,e^{-qw}.$$

Then for $0\leq u<w$,

$$\begin{aligned}|e^{qw}a_c(x,w)-e^{qu}a_c(x,u)|&\leq e^{qw}|a_c(x,w)-a_c(x,u)|+|a_c(x,u)|\,(e^{qu}-e^{qw})\\&\leq e^{qw}[V(x,y)]_u^w+\{|a_c(x,0)|+2M_2(\tfrac{1}{2}q,0)\,e^{-\frac{1}{2}qu}\}\,(e^{qu}-e^{qw})\\&\leq (LS)\int_u^w e^{qy}\,dV_y(x,y)+\{|a_c(x,0)|+2M_2(\tfrac{1}{2}q,0)\}\,(LS)\int_u^w e^{-\frac{1}{2}qy}\,d_y(-e^{qy}).\end{aligned}$$

It follows from (28), since $a_c(x,0)=a(x,0+)$, and taking $q<0$,

$$M_1(q,x)\leq 2M_2(q,0)+2|a(x,0+)|+4M_2(\tfrac{1}{2}q,0),$$

$$\limsup_{x\to\infty}M_1(q,x)\leq 2M_2(q,0)+2+4M_2(\tfrac{1}{2}q,0)<\infty.$$

5. *Proof of the sufficiency in Theorem 2*

We prove first that $B(F;z,x)\to F(z)$ as $x\to\infty$, for z in the half plane of convergence of the integral for $F(z)$. This follows from the sufficiency in [4], Theorem 3, since condition (i) is (28), the first part of (ii) is contained in (24), (iii) is contained in (29), and (iv) in (31). Finally, the second part of (ii) is replaced by (33), the effect of which is to ensure that if the integral for $F(z)$ is convergent at $z=v$ then $B(F;z,x)\to F(z)$ in $R(\operatorname{re} v)$, but not necessarily at $z=v$.

From (24) we obtain (26), and from this and (33) it follows that

(45) *$B(F;z,x)$ is uniformly bounded, as $x\to\infty$, in each compact set contained in the half plane of convergence of the integral for $F(z)$.*

(46) We now show that *if $F(z)$ is in S_ρ then $B(F;z,x)$ is uniformly bounded in each compact set Y contained in $R(\sigma-\rho)$, whenever the integral for $F(z)$ is convergent in $R(\sigma)$*

and this with the convergence of $B(F;z,x)$ to $F(z)$ in $R(\sigma)$ will give the efficiency $(Z;\rho)$ of $a(x,w)$ for each $F(z)$ in S_ρ, by using VITALI's theorem on the convergence of a sequence of uniformly bounded analytic functions.

We can obviously assume throughout that $0<\rho<p$. As $F(z)$ is in S_ρ we have $\sigma\leq 0$. If $\sigma<0$ let Y be a compact set contained in $R(\sigma-\rho)$. Then Y lies at a positive distance from the half plane $\operatorname{re} z\leq\sigma-\rho$, so that there is a q in $0<q<\rho$ such that Y lies in $R(\sigma-q)$. Let us put $\gamma=q-\sigma-\rho$. Then by (9),

$$(IP)\int_0^w e^{-zu}\,ds(u)=(IP)\int_0^w e^{(-\gamma-z)u}\,d_u\{F(-\gamma,u)-F(-\gamma)\}. \tag{47}$$

Also we obtain

$$(L)\int_w^{w+1}|F(-\gamma,u)-F(-\gamma)|\,du=(L)\int_w^{w+1}\left|-e^{\gamma u}s(u)+(L)\int_u^{\infty}-\gamma e^{\gamma y}s(y)\,dy\right|du$$

$$\leq(L)\int_w^{w+1}e^{\gamma u}|s(u)|\,du+\gamma(L)\int_w^{w+1}(L)\int_u^{\infty}e^{\gamma y}|s(y)|\,dy\,du.$$

But

$$(L)\int_w^{w+1}e^{\gamma u}|s(u)|\,du\leq e^{\gamma(w+1)}(L)\int_w^{w+1}|s(u)|\,du=O(e^{(\gamma-r)w}),$$

$$(L)\int_u^{\infty}e^{\gamma y}|s(y)|\,dy=\sum_{n=0}^{\infty}(L)\int_{u+n}^{u+n+1}e^{\gamma y}|s(y)|\,dy=O(e^{(\gamma-r)u}),$$

(48) so that $F(-\gamma,u)-F(-\gamma)$ satisfies (27) with r replaced by $-\gamma+r$. It follows by (47), (48), that it is sufficient to prove (46) for $\sigma=0$.

Let Y be a compact set in $R(-r)$. Then, as in the preceding proof, there is a q in $0<q<r$ such that Y lies in $R(-q)$. Let n be an integer such that

(49) $$n(r-q)>1,$$

and let l be an integer in

(50) $$qn-1<l<rn-1,$$

the existence of such an integer l following from (49). From (27) it follows that for numbers $A>0$, $N>0$,

(51) $$|s(w)|\leq A,\qquad (L)\int_w^{w+1}|s(u)|\,du\leq AN e^{-rw}.$$

Let $W_{t,n}$ be the set of w for which

$$A\exp(-tw/n)\geq|s(w)|>A\exp\{-(t+1)w/n\}.$$

In the set $W_1\equiv\bigcup_{t\geq l+1}W_{t,n}$ we have

$$s(w)=O[\exp\{-(q+\varepsilon)w\}]$$

for some $\varepsilon>0$, since $t/n>q$ by (50). Hence if $s_3(w)\equiv s(w)\,ch(W_1,w)$, the corresponding integral $F(s_3(\cdot);z)\equiv F_2(z)$ will converge in $R(-q)$. Thus by (45) we need only take $t=0,1,2,\ldots,l$, and we have

(52)
$$\begin{cases}
A\exp\{-(t+1)(w+1)/n\}\,\mu\{W_{t,n}\cap(w,w+1)\}\\
\qquad\leq(L)\int_{W_{t,n}\cap(w,w+1)}|s(u)|\,du\leq AN e^{-rw},\\
\mu\{W_{t,n}\cap(w,w+1)\}\leq N\exp\{(t+1)/n\}\exp\{(t+1-rn)w/n\},\\
(LS)\int_0^{\infty}e^{qw}|s(w)-s_3(w)|\,d_wV[a_c(x,w);(0,w)]\\
\leq\sum_{t=0}^{l}(LS)\int_{W_{t,n}}e^{qw}A\exp(-tw/n)\,d_wV[a_c(x,w);(0,w)]\\
\leq\sum_{t=0}^{l}A\,M_2(q-t/n,r-(t+1)/n)\,[1+N\exp\{(t+1)/n\}]<\infty
\end{cases}$$

by (32) for $x \geq \sup_{0 \leq t \leq l} x_0(q - t/n, r - (t+1)/n)$, since $q - t/n < r - (t+1)/n$, which follows from (49). Also, from (17), (34) for $s(0) - s_3(0)$ not necessarily zero, and the definition (2) of IP-integration,

$$(53) \quad \begin{cases} B(F - F_2; z, x) = -(PS)\int_0^\infty \{F(z,w) - F_2(z,w)\}\, d_w a(x,w) \\ \qquad = -(PS)\int_0^\infty [e^{-zw}\{s(w) - s_3(w)\} - \\ \qquad\quad - \{s(0) - s_3(0)\} + z F_3(z,w)]\, d_w a(x,w), \end{cases}$$

where

$$(54) \quad \begin{cases} F_3(z,w) = (IP)\int_0^w e^{-zu}\, d_u \left[(L)\int_0^u \{s(y) - s_3(y)\}\, dy\right] \\ \qquad = (IP)\int_0^w -e^{-zu}\, d_u \left[(L)\int_u^\infty \{s(y) - s_3(y)\}\, dy\right] \\ \qquad = (L)\int_0^\infty \{s(y) - s_3(y)\}\, dy - z(PS)\int_0^w e^{-zu}\left[(L)\int_u^\infty \{s(y) - s_3(y)\}\, dy\right] du. \end{cases}$$

By (51), the last PS-integral is convergent as $w \to \infty$, absolutely and uniformly for z in $R(-q)$. Hence from (54),

$$F_3(z,w) = F_3(z) + O(|z|\, e^{(q-r)w}), \quad \text{where} \quad F_3(z) \equiv \lim_{w \to \infty} F_3(z,w),$$

uniformly for z in $R(-q)$. Substituting in (53),

$$B(F - F_2; z, x) = \{-s(0) + s_3(0) + z F_3(z)\}\, a(x,0) - \\ - (PS)\int_0^\infty [e^{-zw}\{s(w) - s_3(w)\} + O(|z|^2 e^{(q-r)})]\, d_w a(x,w)$$

and (46) follows from (52), (33).

6. *The general class of functions* $F(z)$

If $a(x,w)$ is of bounded variation in w put[10])

$$Q(x,z) \equiv (L)\int_0^\infty e^{-zw}\, a'(x,w)\, dw.$$

Then $a(x,w)$ will be said to *satisfy condition* A_p if there are numbers $q < p$, with limit-point p, and to each such q there are an integer $n \geq 0$, and numbers $1 \leq r \leq 2$, $\nu \geq 0$, $x_0 \geq 0$, all depending on q, such that

$$(55) \quad (R)\int_{|\varrho| \geq \nu} |Q^{(n)}(x, -q + i\varrho)|^r\, d\varrho \leq M_4(q) < \infty$$

[10]) Here we denote $\frac{\partial}{\partial w} a(x,w)$ by $a^1(x,w)$.

for $x \geq x_0$, where $M_4(q)$ is independent of x, and where the differentiation is with respect to $i\varrho$.

Theorem 3. *If $a(x, w)$ is a finite* BAIRE *function satisfying* (6) *and condition A_p, then in order that $a(x, w)$ should be efficient $(Z; p)$ for each $F(z)$ satisfying* (7), (8), *with $s(w)$ a* BAIRE *function bounded in each finite interval of $w \geq 0$, it is necessary and sufficient that* (24) *is true for all $x \geq 0$, $q < p$; that* (28), (29), (31), (32) *are all true; and that*

$Q(x, z)$ tends to 0 as $x \to \infty$, uniformly for z in each compact set contained in

$$R(-p) \cap CZ. \tag{56}$$

The necessity of (24), (28), (29), (31), (32) follows as in Theorem 2. To prove the necessity of (56) we note that in $R(0)$,

$$\frac{1}{z}(1 - e^{-zw}) = (R)\int_0^w e^{-zw}\,dw, \qquad \frac{1}{z} = (R)\int_0^\infty e^{-zw}\,dw,$$

the sole singularity of the analytic function z^{-1} being at $z = 0$. By hypothesis and then (24),

$$(IP)\int_0^\infty a(x, w)\,d_w\left\{\frac{1 - e^{-zw}}{z}\right\} \to \frac{1}{z},$$

$$(IP)\int_0^\infty a(x, w)\,d_w\,e^{-zw} \to -1,$$

$$-a(x, 0) - \sum_{w>0} e^{-zw}\{a(x, w+) - a(x, w-)\} - (LS)\int_0^\infty e^{-zw}\,d_w\,a_c(x, w) \to -1,$$

so that by (28), (29),

$$(LS)\int_0^\infty e^{-zw}\,d_w\,a_c(x, w) \to 0 \tag{57}$$

as $x \to \infty$, uniformly in each compact set in $R(-p) \cap CZ$.

Now let $W(x)$ be the set of measure zero that contains all the variation of

$$a_c(x, w) - (L)\int_0^w a'(x, u)\,du.$$

Take a sequence $\{x_n\}$ of x tending to infinity. Then the set

$$W_2 \equiv \bigcup_{n=1}^{\infty} W(x_n)$$

is of measure zero, and the function $s_4(w) \equiv ch(W_2; w)$ satisfies (27) for all r. Further, for all z in $R(0)$,

$$F_4(z) \equiv F(s_4(\cdot); z) = -s_4(0),$$

since

$$F_4(z, w) = e^{-zw}s_4(w) - s_4(0).$$

Hence by Theorem 2, sufficiency, and (28), (29),

$$(LS)\int_0^\infty e^{-zw}\,d_w\,a_c(x_n,w) - (L)\int_0^\infty e^{-zw}\,a'(x_n,w)\,dw$$

$$= (LS)\int_0^\infty e^{-zw}\,ch(W_2;w)\,d_w\,a_c(x_n,w)$$

$$= (LS)\int_0^\infty \{s_4(0)+F_4(z,w)\}\,d_w\,a_c(x_n,w)$$

$$= -s_4(0)\,a_c(x_n,0) - \sum_{w>0} F_4(z,w)\,\{a(x_n,w+)-a(x_n,w-)\} -$$

$$-(IP)\int_0^\infty a(x_n,w)\,d_w F_4(z,w) \to -s_4(0)+s_4(0) = 0$$

as $n\to\infty$, and is uniformly bounded in each compact set contained in $R(-p)$. This is true for each sequence $\{x_n\}$ tending to infinity, and so is true as $x\to\infty$. Hence (56) is true, using (57).

7. *Proof of the sufficiency in Theorem 3*

Let Y be a compact set in $R(\sigma-p)$. Then we can find $q<p$ such that for all z in Y, $\operatorname{re} z\geq\sigma-p$. Taking $v=\sigma+\varepsilon$ in (10) we obtain

$$F(z,w) = O\{e^{(q+\varepsilon)w}\} \tag{58}$$

as $w\to\infty$, uniformly for z in Y, for each $\varepsilon>0$. There is an $\varepsilon>0$ such that $q+2\varepsilon<p$, so that from (24),

$$|a(x,w)\,F(z,w)| = O\{e^{-(q+2\varepsilon)w}e^{(q+\varepsilon)w}\} = o(1)$$

as $w\to\infty$. By Theorem 1 the integral for $B(F;z,x)$ is convergent in Y. Thus

$$\begin{cases} B(F;z,x) = -(PS)\int_0^\infty F(z,w)\,d_w a(x,w) \\ = -\sum_{w>0} F(z,w)\,\{a(x,w+)-a(x,w-)\} - (LS)\int_0^\infty F(z,w)\,d_w a_c(x,w). \end{cases} \tag{59}$$

Using (29) and (58), the first term is uniformly bounded in Y and tends to 0 there as $x\to\infty$. Further, following a proof similar to that at the end of section 6, we can easily show that

$$(LS)\int_0^\infty F(z,w)\,d_w a_c(x,w) - (L)\int_0^\infty F(z,w)\,a'(x,w)\,dw \to 0$$

as $x\to\infty$, and is uniformly bounded in Y. Hence from (59),

$$\begin{cases} B(F;z,x) = o(1) - (L)\int_0^\infty F(z,w)\,a'(x,w)\,dw \\ \equiv o(1) + B^*(F;z,x), \end{cases} \tag{60}$$

as $x\to\infty$, the $o(1)$ being uniformly bounded in Y, where $B^*(F;z,x)$ is the $B(F;z,x)$ obtained by replacing $a(x,w)$ by

$$a^*(x,w) \equiv a(x,0+) + (L)\int_0^w a'(x,u)\,du.$$

We can also replace $s(w)$ by any other BAIRE function $s_\beta(w)$, provided that $s(w)-s_\beta(w)$ satisfies (27) for $\tau=\beta$. In particular, we can put

$$s(w)=0 \quad \text{for} \quad 0\leq w\leq 1. \tag{61}$$

The condition (61) will be useful later. Let us consider the integral

$$G(z,x,w)\equiv\frac{1}{2\pi i}(R)\int_{-w}^{w}\frac{F(z+i\varrho+q)\,Q(x,-i\varrho-q)}{\varrho-iq}\,d\varrho. \tag{62}$$

To ensure that $z+i\varrho+q$ is in the region of convergence of the LAPLACE-STIELTJES integral for $F(z)$, we take

$$\operatorname{re}(z+i\varrho+q)>\sigma, \quad \operatorname{re} z>\sigma-q,$$

since ϱ is real. As $q<\beta$, and as q is arbitrarily near to β, we can deal with all points of the region $R(\sigma-\beta)$, but, in general, with no points outside that region. From (10), with $v=\operatorname{re} z+q-\varepsilon$, where $\varepsilon=\frac{1}{2}(\operatorname{re} z+q-\sigma)>0$, we obtain (58) with $q+\varepsilon$ replaced by $q-\varepsilon$. Hence from $(11)_\beta$

$$F(z+i\varrho+q)=(i\varrho+q)\,(L)\int_0^\infty e^{-(i\varrho+q)w}F(z,w)\,dw$$

the integral being absolutely convergent, and $e^{-qw}F(z,w)$ is bounded and $L^2(0,\infty)$.

Also let $q+\delta<\beta$. Then by (24), and Lemma 2 for $q+\delta$,

$$e^{qw}|a'(x,w)|\leq q\,e^{qw}|a(x,w)|+|q\,e^{qw}a(x,w)+e^{qw}a'(x,w)|,$$

$$(L)\int_0^\infty e^{qw}|a'(x,w)|\,dw\leq q\,\delta^{-1}M_1(q+\delta,x)+M_1(q,x)<\infty,$$

and $e^{qw}a'(x,w)$ is in $L^1(0,\infty)$, for each $x\geq 0$, $q<\beta$. Hence applying a PARSEVAL Theorem, TITCHMARSH [8], Theorem 70, p. 94, to the functions $e^{qw}a'(x,w)$ and $e^{qw}F(z,-w)$, where $a'(x,w)$ and $F(z,w)$ are defined to be zero in $w<0$, we see that

$$(C,1)\lim_{w\to\infty}G(z,x,w)\equiv\lim_{w\to\infty}\frac{1}{2\pi i}(R)\int_{-w}^{w}\left(1-\frac{|\varrho|}{w}\right)\frac{F(z+i\varrho+q)\,Q(x,-i\varrho-q)}{\varrho-iq}\,d\varrho$$

$$=(L)\int_0^\infty F(z,w)\,a'(x,w)\,dw=-B^*(F;z,x).$$

We now put $-v=i\varrho+q$, so that

$$B^*(F;z,x)=\frac{1}{2\pi i}(R)\int_\vartheta\frac{F(z-v)\,Q(x,v)}{v}\,dv\,(C,1) \tag{63}$$

for the path ϑ given by $\operatorname{re} v=-q$, $-\infty<\operatorname{im} v<\infty$.

By (13), (14), (15), (16), we can displace a finite part of the path ϑ to the right, so that it surrounds the set Z and passes to the right of the origin,

and yet is unaltered in $|\operatorname{im} v| \geq \delta$ for some $\delta > 0$. Let us denote the new path by φ, and let $\varphi(w)$ be the finite part of φ that terminates at $v = -q \pm iw$ $(w \geq \delta)$.

From (12), Z is closed, and also the set of points $\alpha + i\beta$ of $F(z)$ is closed, so that $CH(F; Z)$ is closed. Let Y be a compact set lying in

$$H(F; Z) \cap R(\sigma - p).$$

Then Y lies at a positive distance Δ from the sum of $CH(F; Z)$ and the half plane given by $\operatorname{re} v \leq \sigma - p$, so that if each point of $\varphi(\delta)$ is nearer than $\Delta/2$ to Z, we can use the same path φ for each point z of Y. In this case $F(z-v)$ is bounded uniformly for z in Y and v on $\varphi(w)$.

Also $Q(x, v) \to 0$ as $x \to \infty$, uniformly for v on $\varphi(w)$, though not necessarily for v on φ, by (56).

Displacing the path from ϑ to φ we pass over a pole at $v = 0$, the residue of the integrand being $F(z)\, Q(x, 0)$. Since

$$Q(x, 0) = -a^*(x, 0) = -a(x, 0+) \to -1$$

as $x \to \infty$, the residue tends to $-F(z)$. Thus by (63),

$$(64)\quad \begin{cases} B^*(F; z, x) = F(z)\, a(x, 0+) - \dfrac{1}{2\pi i}\,(R)\displaystyle\int_{\varphi(w)} \frac{F(z-v)\, Q(x, v)}{v}\, dv - \\ \qquad - \lim\limits_{u \to \infty} \dfrac{1}{2\pi i}\,(R)\displaystyle\int_{\varphi(u)-\varphi(w)} \left(1 - \frac{|\operatorname{im} v|}{u}\right) \frac{F(z-v)\, Q(x, v)}{v}\, dv, \\ \lim\limits_{x \to \infty} B^*(F; z, x) \\ \qquad = F(z) - \lim\limits_{x \to \infty} \lim\limits_{u \to \infty} \dfrac{1}{2\pi i}\,(R)\displaystyle\int_{\varphi(u)-\varphi(w)} \left(1 - \frac{|\operatorname{im} v|}{u}\right) \frac{F(z-v)\, Q(x, v)}{v}\, dv, \end{cases}$$

uniformly for z in Y. Thus by (60), to show that $a(x, w)$ is efficient $(Z; p)$ for $F(z)$ it is sufficient to show that the last double limit is 0, uniformly for z in Y. This is where we use condition A_p'.

By (11),

$$(R)\int_{-q-i\varrho}^{-q+iu} \frac{F(z-\zeta)}{\zeta}\, d\zeta = (R)\int_{-q-i\varrho}^{-q+iu} \left\{ -(L)\int_0^\infty e^{\zeta w} F(z, w)\, dw \right\} d\zeta$$

$$= (L)\int_0^\infty \frac{F(z, w)}{w}\, e^{-qw} \left(e^{-i\varrho w} - e^{iuw}\right) dw.$$

The inversion is possible since the double integral is absolutely convergent. By (61),

$$(65)\qquad F(z, w) = 0 \quad \text{for} \quad 0 \leq w \leq 1,$$

so that we can take $\varrho \to \infty$ and use the RIEMANN-LEBESGUE Theorem, obtaining

$$(66)\qquad (R)\int_{-q-i\infty}^{-q+iu} \frac{F(z-\zeta)}{\zeta}\, d\zeta = -(L)\int_0^\infty \frac{F(z, w)}{w}\, e^{(-q+iu)w}\, dw.$$

Thus we can obviously integrate n times between $-q-i\infty$ and $-q+iu$, in which case $w^{-1}F(z,w)$ in (66) is replaced by $w^{-n}F(z,w)$. Integrating by parts n times, we obtain

$$(67)\quad \begin{cases} (R)\displaystyle\int_{-q+iw}^{-q+iu} \frac{F(z-v)}{v}\,Q(x,v)\,dv = [Q(x,v)\,F_{(1)}(z,v) - Q'(x,v)\,F_{(2)}(z,v) + \cdots \\ \qquad\qquad + (-)^{n-1}\,Q^{(n-1)}(x,v)\,F_{(n)}(z,v)]_{v=-q+iw}^{v=-q+iu} \\ \qquad\qquad + (-)^n\,(R)\displaystyle\int_{-q+iw}^{-q+iu} Q^{(n)}(x,v)\,F_{(n)}(z,v)\,dv, \end{cases}$$

where

$$F_{(0)}(z,v) \equiv \frac{F(z-v)}{v}, \qquad F_{(m)}(z,v) \equiv (R)\int_{-q-i\infty}^{v} F_{(m-1)}(z,\zeta)\,d\zeta \qquad (m \geq 1).$$

$$\lim_{u\to\infty} F_{(m)}(z,-q+iu) = 0$$

for $m = 1, 2, \ldots, n$, by the RIEMANN-LEBESGUE Theorem. Also, as

$$e^{(q+\varepsilon)w}\,|a'(x,w)|$$

is in $L^1(0,\infty)$ for $q+\varepsilon<p$,

$$|Q^{(m)}(x,-q+iu)| \leq (L)\int_0^\infty w^m e^{qw}\,|a'(x,w)|\,dw$$

$$= O\Big((L)\int_0^\infty e^{(q+\varepsilon)w}|a'(x,w)|\,dw\Big) = O(1)$$

for $m = 0, 1, \ldots, n-1$, as $u\to\infty$. Hence from (67) we obtain

$$(68)\quad \begin{cases} (R)\displaystyle\int_{-q+iw}^{-q+i\infty} \frac{F(z-v)}{v}\,Q(x,v)\,dv = -\,Q(x,-q+iw)\,F_{(1)}(z,-q+iw) + \cdots \\ \qquad\qquad + (-)^n\,Q^{(n-1)}(x,-q+iw)\,F_{(n)}(z,-q+iw) + \\ \qquad\qquad + (-)^n\,(R)\displaystyle\int_{-q+iw}^{-q+i\infty} Q^{(n)}(x,v)\,F_{(n)}(z,v)\,dv. \end{cases}$$

Now by (56) and CAUCHY's integrals we see that for fixed $w>0$,

$$(69)\qquad Q^{(m)}(x,-q+iw)\to 0 \quad \text{as} \quad x\to\infty \qquad (0\leq m<n).$$

Also, by (55), we see that $Q^{(n)}(x,-q+i\varrho)$ is in $L^r(\nu,\infty)$ and in

$$(70)\qquad L^r(-\infty,-\nu).$$

From (58), with the q there taken to be greater than $\sigma-\operatorname{re} z$, and from (65), we see that

$$\sup_{x\in Y}(L)\int_0^\infty |w^{-n}F(z,w)\,e^{-kw}|^r\,dw < \infty \qquad (r>0,\ k>\sigma-\operatorname{re} z).$$

Hence if $1<r\leq 2$ and if r' is defined by $r+r'=rr'$, we have by [8], p. 96, Theorem 74, (4.1.2),

$$\sup_{z\in Y}(R)\int_{-\infty}^{\infty}|F_{(n)}(z,-k+iu)|^{r'}du<\infty \qquad (k>\sigma-\operatorname{re} z).$$

Hence putting $k=q<p$, and also using (68), (69), (70), we obtain, for $w\geq r$,

$$\limsup_{x\to\infty}\left|(R)\int_{-q+iw}^{-q+i\infty}\frac{F(z-v)}{v}Q(x,v)\,dv\right| = \limsup_{x\to\infty}\left|(R)\int_{-q+iw}^{-q+i\infty}Q^{(n)}(x,v)\,F_{(n)}(z,v)\,dv\right|$$

$$\leq \limsup_{x\to\infty}\left\{(R)\int_{w}^{\infty}|Q^{(n)}(x,-q+i\varrho)|^{r}\,d\varrho\right\}^{\frac{1}{r}}\left\{(R)\int_{w}^{\infty}|F_{(n)}(z,-q+i\varrho)|^{r'}\,d\varrho\right\}^{\frac{1}{r'}}$$

$$\leq M_4^{1/r}(q)\left\{(R)\int_{w}^{\infty}|F_{(n)}(z,-q+i\varrho)|^{r'}\,d\varrho\right\}^{\frac{1}{r'}}.$$

Similarly for the range $-q-i\infty$ to $-q-iw$.

For $r=1$ we note that $F_{(n)}(z,-q+i\varrho)$ is bounded in ϱ, uniformly for z in Y. A similar result follows in this case.

From (64) it follows that

$$B^*(F;z,x)\to F(z) \text{ as } x\to\infty, \text{ uniformly for } z \text{ in } Y.$$

8. Examples of functions $Q(x,z)$, and the connection with OKADA's *theory*

If we transform the complex plane minus the line $v=\operatorname{re} v\geq 1$ into the plane minus the line $z=\operatorname{re} z\leq 0$, and put

$$Q(x,z)\,E(x)=-E(x,v),$$

where $E(v)$ is an integral function occurring in COOKE [1], Theorem (8.2, III), p. 184, Theorem (8.3, I), p. 188, it seems likely that some of these functions $Q(x,z)$ will satisfy the conditions of Theorem 3. Two such transformations are $v=e^{-z}$, $v=1-z$.

Alternatively, we can transform CZ into $|v|<1$, and $Q(x,z)$ into $Q_1(x,v)$, the latter tending to 0 as $x\to\infty$, uniformly in each circle $|v|\leq 1-\varepsilon$ for $\varepsilon>0$. Theoretically, the search for such functions $Q(x,z)$ is thereby simplified. A simple example is the following.

If $0<\varepsilon<1$, $-v^x\to 0$ uniformly in $|v|\leq 1-\varepsilon$ as $x\to\infty$. Putting $v=(1+z)^{-1}$, so that $|v|<1$ becomes $|z+1|>1$, we obtain

$$Q(x,z)=-(1+z)^{-x}=-\Gamma^{-1}(x)\,(R)\int_{0}^{\infty}e^{-zu}e^{-u}u^{x-1}\,du,$$

and a suitable function $a(x,w)$ is given by

$$a(x,w)=\Gamma^{-1}(x)\,(R)\int_{w}^{\infty}e^{-u}u^{x-1}\,du,$$

$$a(n,w)=\left\{1+w+\frac{w^2}{2!}+\cdots+\frac{w^{n-1}}{(n-1)!}\right\}e^{-w}$$

for integers $n \geq 1$. The integral for $Q(x, z)$ is convergent in $R(-1)$, so that $p = 1$. Also

$$|Q(x, z)| = \{(\operatorname{re} z + 1)^2 + (\operatorname{im} z)^2\}^{-\frac{1}{2}x}$$

so that Q satisfies condition A_1 with

$$q < 1, \quad n = 0, \quad r = 1, \quad x_0 = 2, \quad \nu = 1.$$

Taking Z as including the region given by $|z+1| \leq 1$, it is easy to verify that the conditions of Theorem 3 are satisfied. Thus $a(x, w)$ is efficient $(Z; 1)$ for all $F(z)$ given by (7), (8), the value of $a(n, w)$ being particularly simple, and we obtain

$$B(F; z, n) = F(z+1) - F'(z+1) + \frac{F''(z+1)}{2!} - \cdots + (-)^{n-1} \frac{F^{(n-1)}(z+1)}{(n-1)!},$$

the n-th partial sum of a TAYLOR series.

Further examples can be obtained in the following way.

Theorem 4. *Let $\alpha > 0$, $\beta > 0$, $\xi > 0$, $\eta > 0$, with $-\alpha - i\xi$, $-\beta + i\eta$ in CZ, where Z is an H-set, and let Z_1 be the strip*

$$-\xi \leq \operatorname{im} z \leq \eta, \quad \operatorname{re} z \geq -\max(\alpha, \beta).$$

For each x let $J(x, z)$ be regular for z in Z_1, apart from a simple pole at $z = 0$, with residue tending to -1 as $x \to \infty$. Further, let $J(x, z) \to 0$ as $x \to \infty$, uniformly for each compact set contained in $Z_1 - Z$. If $F(z)$ is convergent in $R(\sigma)$ then

$$(71) \quad F(z) = \lim_{x \to \infty} \frac{1}{2\pi i} (R) \int_\gamma F(z - v)\, J(x, v)\, dv = \lim_{x \to \infty} (IP) \int_0^\infty a(x, w)\, d_w F(z, w),$$

uniformly for z in each compact set contained in

$$H(F; Z) \cap R(\sigma - \min(\alpha, \beta)),$$

where γ is the line from $-\alpha - i\xi$ to $-\beta + i\eta$, and where

$$(72) \qquad a(x, w) \equiv \frac{1}{2\pi i} (R) \int_\gamma e^{vw} J(x, v)\, dv.$$

Thus the function given by (72) is efficient $(Z; p)$ for all $F(z)$ satisfying (7), (8), where $p = \min(\alpha, \beta)$.

To prove the efficiency $(Z; p)$ we deform γ to a new path in the same way as part of ϑ was deformed to become $\varphi(\delta)$. Further, by (11) and the absolute convergence of the integrals, if z lies in $R(\sigma - \min(\alpha, \beta))$,

$$\begin{aligned}
(R) \int_\gamma F(z - v)\, J(x, v)\, dv &= -(R) \int_\gamma \left\{(L) \int_0^\infty v\, e^{vw} F(z, w)\, dw\right\} J(x, v)\, dv \\
&= (L) \int_0^\infty F(z, w) \left\{-(R) \int_\gamma v\, e^{vw} J(x, v)\, dv\right\} dw \\
&= (LS) \int_0^\infty F(z, w)\, d_w \left\{-(R) \int_\gamma e^{vw} J(x, v)\, dv\right\} \\
&= (IP) \int_0^\infty \left\{(R) \int_\gamma e^{vw} J(x, v)\, dv\right\} dF(z, w),
\end{aligned}$$

giving the second part of (71) if $a(x, w)$ is given by (72). This $a(x, w)$ is, in particular, efficient $(Z; p)$ for functions $F(z)$ of class S_r, for all r, so that by Theorem 2, the conditions (24), (28), (29), (31), (32) are satisfied. Further, $a(x, w)$ satisfies condition A_p, since

$$2\pi i\, Q(x, z) = (L)\int_0^\infty e^{-zw}\left\{(R)\int_\gamma v\, e^{vw} J(x, v)\, dv\right\} dw$$

$$= (R)\int_\gamma v\, J(x, v)\left\{(R)\int_0^\infty e^{(v-z)w}\, dw\right\} dv$$

$$= (R)\int_\gamma \frac{v}{z - v} J(x, v)\, dv,$$

if $\operatorname{re}(v - z) < 0$ for v on γ, and so $\operatorname{re} z > -\min(\alpha, \beta)$. As $J(x, v)$ is uniformly convergent on γ as $x \to \infty$, then $J(x, v)$ is uniformly bounded for v on γ and $x \geq x_0$, for some x_0. Thus

$$Q(x, z) = O(\max |z - v|^{-1}) = O(|z|^{-1}) \quad \text{as} \quad |\operatorname{im} z| \to \infty,$$

and satisfies condition A_p with $n = 0$, $x \geq x_0$, $r = 2$, ν large. Thus it follows that Theorem 4 is a special case of Theorem 3, sufficiency.

It can easily be proved from (72) that $a(x, w)$ *is an integral function of order* < 1, *or of order* 1, *type*

$$\leq \max(\alpha, \beta) + \max(\xi, \eta). \tag{73}$$

Specializing still further, let us take

$$\alpha = \beta = p,\ \xi = \eta = \pi,\ J(x, z) \equiv (1 - e^z)^{-1} K(x, e^{-z}),\ K(x, v) \equiv \sum_{n=0}^{\infty} k_n(x)\, v^n. \tag{74}$$

Let Y_1 denote the region $|v| \leq e^p$, and let Y be the set of $v = e^{-z}$ for which z lies in Z. Then this special $J(x, z)$ satisfies the conditions of Theorem 4 if and only if

$$K(x, v) \text{ is convergent in } |v| \leq e^p, \tag{75}$$

$$K(x, 1) \to 1 \text{ as } x \to \infty, \text{ and} \tag{76}$$

$$K(x, v) \to 0 \text{ as } x \to \infty, \text{ uniformly for each compact set contained in } Y_1 - Y. \tag{77}$$

Further, in this case, for integers n,

$$a(x, n) = \sum_{m=n}^{\infty} k_m(x),\ k_n(x) = a(x, n) - a(x, n + 1). \tag{78}$$

This special case corresponds to the theory of [6], p. 68, § 3, Satz III, OKADA's summation of TAYLOR series by convergence factors. *Conditions* (75), (76), (77) *are necessary and sufficient in order that*

$$\sum_{n=0}^{\infty} a(x, n)\, f_n v^n$$

should converge in $|v| \leq e^p r$, *and should tend to* $f(v)$ *as* $x \to \infty$, *uniformly in each compact set contained in that part of the partial star domain of* $f(z)$ *cor-*

responding to Y that is contained in $|v| \leq e^{\rho} r$, for all

$$f(v) \equiv \sum_{n=0}^{\infty} f_n v^n,$$

convergent in $|v| \leq r$. Thus by (78), OKADA's Theorem is a special case of Theorem 4.

Further, from each convergence factor $a(x, n)$ satisfying OKADA's Theorem we can put

$$k_n(x) = a(x, n) - a(x, n+1)$$

and so construct $J(x, z)$ as in (74). Substituting in (72) we obtain a convergence factor $a(x, w)$ depending on two continuous parameters $x \geq 0$, $w \geq 0$, and coinciding with $a(x, n)$ for integer values n of w, and satisfying the conditions of Theorem 3. Thus Theorem 4 is an important special case of Theorem 3, sufficiency. However, (73) shows that the convergence factors produced by (72) are of a very special kind.

Finally, we may note difficulties in the weakening or removal of condition A_ρ in Theorem 3. We know by (56) that $Q(x, v)$ tends to 0 as $x \to \infty$, uniformly for v on any finite part of the path φ. But without an extra condition on the product

$$(\varrho - iq)^{-1} F(z + i\varrho + q)\, Q(x, -i\varrho - q)$$

as a function of ϱ we cannot prove that the integral over $|\varrho| \geq w$ tends to 0 as $x \to \infty$, for fixed $w \geq \delta$. The extra condition is supplied by (55). The value of r there cannot in general be taken greater than 2. For although

$$w^{-n} F(z, w)\, e^{-qw} = O(e^{-\varepsilon w})$$

for some $\varepsilon > 0$, we cannot show in general that the function $F_{(n)}(z, -q + iu)$ of u is in $L^k(-\infty, \infty)$ for $k < 2$.

References

[1] COOKE, R. G.: Infinite matrices and sequence spaces. London: Macmillan 1950. — [2] HALMOS, P. R.: Measure theory. New York 1950. — [3] HENSTOCK, R.: The efficiency of matrices for Taylor series. J. London Math. Soc. **22**, 104—107 (1947). — [4] HENSTOCK, R.: The efficiency of convergence factors for functions of a continuous real variable. J. London Math. Soc. **30**, 273—286 (1955). — [5] HENSTOCK, R.: On Ward's Perron-Stieltjes integral. To be published, Canadian J. of Mathematics. — [6] OKADA, Y.: Über die Annäherung analytischer Funktionen. Math. Z. **23**, 62—71 (1925). — [7] SAKS, S.: Theory of the integral, 2nd English ed. Warsaw 1937. — [8] TITCHMARSH, E. C.: Introduction to the theory of Fourier integrals, 2nd ed. Oxford 1948. — [9] WARD, A. J.: The Perron-Stieltjes integral. Math. Z. **41**, 578—604 (1936). — [10] ZAANEN, A. C.: Linear analysis. Amsterdam 1953.

Department of Pure Mathematics, The Queen's University, Belfast (Nord-Irland)

(Eingegangen am 23. Mai 1956)

WALTER, W.
Math. Zeitschr. Bd. 67, S. 32—37 (1957)

Ganze Lösungen der Differentialgleichung $\Delta^p u = f(u)$

Von
WOLFGANG WALTER

1. Es ist bekannt, daß die Differentialgleichung $\Delta u = f(u)$ keine ganzen Lösungen besitzt, wenn $f(s)$ stetig und positiv ist und für große s stärker wächst als $s^{1+\alpha}$ mit $\alpha > 0$ [2], [3], [5][1]). Das Ziel der vorliegenden Arbeit ist, den analogen Satz für die Differentialgleichung

$$\Delta^p u = f(u) \quad (p \geq 2) \tag{1}$$

zu beweisen. Der Operator Δ^p ist wie üblich definiert durch

$$\Delta = \sum_{\nu=1}^{n} \frac{\partial^2}{\partial x_\nu^2}, \qquad \Delta^p = \Delta(\Delta^{p-1}).$$

Unter einer ganzen Lösung verstehen wir eine im ganzen n-dimensionalen Euklidischen Raum R_n $2p$-mal stetig differenzierbare Lösung $u(x) = u(x_1, \ldots, x_n)$ der vorgelegten Differentialgleichung.

Es wird sich erweisen, daß man ohne weitere Voraussetzungen über $f(s)$ nur im Falle $n = 2$ zum Ziele kommt. Bei mehr als zwei unabhängigen Veränderlichen muß man darüber hinaus $f(s) \geq \delta > 0$ voraussetzen. Das ist nicht nur ein Mangel der Beweismethode; vielmehr läßt sich zeigen, daß diese zusätzliche Bedingung wesentlich für die Gültigkeit des Satzes ist. Sie tritt auch bei HAVILAND [2], der den Fall $p = 1$ behandelt, auf; doch ist sie da, wie der Verfasser [3] gezeigt hat, entbehrlich.

2. Satz. *Die Funktion $f(s)$ sei für $-\infty < s < \infty$ stetig und positiv; es gebe ferner zwei positive Konstanten α und s_0, so daß*

$$f(s) \geq s^{1+\alpha} \quad \text{für} \quad s \geq s_0 \tag{2}$$

ist.

Dann gibt es im R_2 keine ganze Lösung $u(x_1, x_2)$ der Differentialgleichung (1). Ist außerdem $f(s) \geq \delta > 0$, so existiert auch für $n > 2$ keine ganze Lösung $u(x_1, \ldots, x_n)$. Ohne diese zusätzliche Voraussetzung wird die Behauptung im allgemeinen falsch.

Beweis. Entgegen der Behauptung sei $u(x)$ eine ganze Lösung von (1) mit $p \geq 2$. Die Funktion $v = \Delta^{p-1} u$ ist nicht nach oben beschränkt; das folgt leicht aus $\Delta v > 0$ $(n = 2)$ bzw. $\Delta v \geq \delta > 0$ $(n > 2)$[2]). Durch eine lineare Trans-

[1]) Zahlen in eckigen Klammern verweisen auf das Literaturverzeichnis.

[2]) Für $n = 2$ ist der Mittelwert von v bekanntlich eine konvexe Funktion von $\log R$, für $n > 2$ wächst er unter den angegebenen Voraussetzungen sogar wie R^2.

formation können wir also erreichen, daß $v(0) = \Delta^{p-1} u(0) > 0$ ist. Der Mittelwert von u für die Kugeloberfläche vom Radius R

$$m(R, u) = \frac{1}{\omega R^{n-1}} \oint_{|x|=R} u(x)\, dx \qquad (\omega = 2\pi^{n/2}/\Gamma(n/2))$$

läßt sich in eine Reihe entwickeln (s. [4] § 1):

$$(3) \qquad m(R, u) = \sum_{i=0}^{p-1} a_{n,i} R^{2i} \Delta^i u(0) + \int_0^R m(r; \Delta^p u)\, r^{n-1} H_p(r, R)\, dr,$$

wobei

$$a_{n,i} = \frac{1}{2^i i!\,(n+2)\dots(n+2i-2)}$$

und

$$(4) \quad H_p(r, R) = A_p \int_r^R [(R^2 - s^2)(s^2 - r^2)]^{p-1} s^{3-n-2p}\, ds, \qquad A_p = 4^{1-p}(p-1)!^{-2}$$

ist. Die folgenden Rechnungen vereinfachen sich, wenn wir statt $m(R, u)$ die Funktion $M(T, u) = m(\sqrt{T}, u)$ einführen. Aus (3) wird dann, wie die Transformation $T = R^2$, $t = r^2$ zeigt,

$$(5) \quad M(T, u) = \sum_{i=0}^{p-1} a_{n,i} T^i \Delta^i u(0) + \int_0^T M(t, \Delta^p u)\, t^{m-1} H_p'(t, T)\, dt \quad (m = n/2)$$

mit

$$(6) \qquad H_p'(t, T) = A_p' \int_t^T [(T - z)(z - t)]^{p-1} z^{1-m-p}\, dz, \qquad A_p' = A_p/4.$$

Die Ableitungen von $M(T, u)$ ergeben sich wegen

$$(7) \quad \begin{cases} D^k H_p'(t, T) = A_p'(p-1)\cdots(p-k) \displaystyle\int_t^T (T - z)^{p-k-1} (z - t)^{p-1} z^{1-m-p}\, dz \\ \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad (D^k = d^k/dT^k) \end{cases}$$

zu

$$(8) \quad D^k M(T, u) = D^k \sum_{i=0}^{p-1} a_{n,i} T^i \Delta^i u(0) + \int_0^T M(t, \Delta^p u)\, t^{m-1} D^k H_p'(t, T)\, dt$$

für $0 \le k \le p-1$. Insbesondere ist, da $\Delta^{p-1} u(0)$ nach Voraussetzung positiv ist,

$$(9) \quad D^{p-1} M(T, u) \ge (p-1)!\, a_{n,p-1} \Delta^{p-1} u(0), \qquad \text{also} \qquad \lim_{T\to\infty} M(T, u) = \infty.$$

Wir betrachten den Ausdruck

$$(10) \qquad J(t, T, m, p) = T^p t^{m-1} (T - t)^{1-2p} \int_t^T [(T - z)(z - t)]^{p-1} z^{1-m-p}\, dz$$

für $0 < t \le T$, $p \ge 2$, $2m = n \ge 2$. Für $m = p + 1 + i$, $i = 0, 1, \dots$ läßt sich das Integral explizit angeben; es ist (s. [1] Formel 421.5)

$$(11) \qquad J(t, T, p+1+i, p) = \sum_{j=0}^{i} \binom{i}{j} (t/T)^j B(p+j, p+i-j) \ge B(p, p+i) > 0,$$

wo $B(p, q) = \Gamma(p)\,\Gamma(q)/\Gamma(p+q)$ die Betafunktion bedeutet. Ferner ist

$$(12)\quad \frac{\partial J(t, T, m, p)}{\partial m} = T^p t^{m-1}(T-t)^{1-2p}\int_t^T [(T-z)(z-t)]^{p-1} z^{1-m-p} \log t/z\, dz \leq 0.$$

Aus (11) und (12) folgt, daß (für festes m und p) $J(t, T, m, p)$ eine positive untere Schranke hat, d.h. nach (10), daß

$$(13)\quad (T-t)^{2p-1} \leq c\, T^p t^{m-1} H_p'(t, T) \quad \text{für } 0 < t \leq T \qquad (c = c(m, p) > 0)$$

ist.

Zu $f(s)$ gibt es eine für alle s positive, monoton wachsende, konvexe Funktion $g(s) \leq f(s)$, welche für große s mit der Funktion $s^{1+\alpha}$ übereinstimmt (vgl. [3] S. 96/97). Es sei $t_1 > 0$ so gewählt, daß $\sum_{i=0}^{p-1} a_{n,i} T^i \Delta^i u(0) > 0$ für $T > t_1$ ist. Dann hat die Funktion

$$(14)\quad N(T) = \int_{t_1}^T (T-t)^{2p-1} g\left(M(t, u)\right) dt$$

die Eigenschaften

$$(15)\quad D^i N(T) > 0 \text{ für } 0 \leq i \leq 2p,\ T > t_1;\quad D^{2p} N(T) = (2p-1)!\, g\left(M(T, u)\right).$$

Aus den Eigenschaften von g ergibt sich (vgl. etwa [3] S. 96)

$$g\left(M(T, u)\right) \leq M\left(T, g(u)\right) \leq M\left(T, f(u)\right) = M(T, \Delta^p u)$$

und hieraus nach (5), (13) und (14)

$$(16)\quad N(T) \leq c\, T^p M(T, u).$$

Es ist also

$$D^{2p} N(T) = (2p-1)!\, g\left(M(T, u)\right) \geq (2p-1)!\, g\left(N(T)/c\, T^p\right).$$

Nach den Ungleichungen (15) wächst $N(T)$ stärker als T^p; für große T dürfen wir demnach $g(s)$ durch $s^{1+\alpha}$ ersetzen und erhalten so

$$(17)\quad C\, T^{p+\alpha p} D^{2p} N(T) \geq N(T)^{1+\alpha} \quad \text{mit} \quad C = c^{1+\alpha}/(2p-1)!.$$

Hierbei kann o. B. d. A. $0 < \alpha < 1$ angenommen werden. Die Funktion $N_1(T) = N(e^T)$ genügt dann wegen

$$D^{2p} N_1(T) \geq e^{2pT} N^{(2p)}(e^T)$$

von einer Stelle an, etwa für $T \geq t_2$, der Ungleichung

$$(18)\quad D^{2p} N_1(T) \geq N_1(T)^{1+\alpha}.$$

Die beiden folgenden Hilfssätze werden zeigen, daß es eine solche Funktion $N_1(T)$ nicht geben kann.

Hilfssatz 1. *Es seien zwei für $t \geq t_0$ k-mal ($k \geq 2$) stetig differenzierbare positive Funktionen $y(t)$ und $z(t)$ gegeben, für welche*

$$(19)\quad y^{(k)} = y^{1+\alpha};\quad z^{(k)} \geq z^{1+\alpha} \qquad (\alpha > 0)$$

und

$$0 < y^{(i)}(t_0) < z^{(i)}(t_0) \qquad \text{für } i = 0, \ldots, k$$

ist. Dann ist

$$y(t) < z(t) \qquad \text{für } t \geq t_0.$$

Beweis. Wir bemerken zunächst, daß y und z monoton wachsend gegen ∞ streben. Es ist zu zeigen, daß $w(t) = z(t) - y(t)$ positiv ist. An der Stelle t_0 ist das nach Voraussetzung der Fall. Angenommen die Behauptung sei falsch und t_1 sei die erste Nullstelle von $w(t)$. Dann ist für $t_0 < t < t_1$

$$w^{(k)}(t) = z^{(k)}(t) - y^{(k)}(t) \geq z^{1+\alpha} - y^{1+\alpha} > 0.$$

Deshalb sind auch alle Ableitungen niedrigerer Ordnung von $w(t)$ positiv in (t_0, t_1). Insbesondere ist $w(t)$ monoton wachsend im Widerspruch zur Annahme $w(t_1) = 0$.

Hilfssatz 2. *Ist $z(t)$ k-mal stetig differenzierbar und*

$$z^{(k)}(t) \geq z^{1+\alpha}(t) \qquad \text{für } t \geq t_0, \quad z^{(i)}(t_0) > 0 \qquad \text{für } i = 0, \ldots, k,$$

so gilt $\lim\limits_{t \to t_1} z(t) = \infty$ *für ein passend gewähltes* $t_1 < \infty$.

Beweis. Die Funktion

$$y(t) = b(t^* - t)^{-\frac{k}{\alpha}}$$

ist, wenn b der Gleichung $b^\alpha = k(k+\alpha) \ldots (k+(k-1)\alpha)\alpha^{-k}$ genügt, eine Lösung der Differentialgleichung

$$y^{(k)} = y^{1+\alpha}.$$

Wählen wir die Konstante t^* $(> t_0)$ genügend groß, so können wir erreichen daß $0 < y^{(i)}(t_0) < z^{(i)}(t_0)$, also nach Hilfssatz 1 $z(t) > y(t)$ für $t > t_0$ ist. Wegen $\lim\limits_{t \to t^*} y(t) = \infty$ existiert, wie behauptet, eine Zahl $t_1 \leq t^*$ mit der angegebenen Eigenschaft.

Wegen der Ungleichungen (15) kann der Hilfssatz 2 auf $N_1(T)$ angewandt werden (mit $k = 2p$). Damit ist gezeigt, daß $N_1(T)$ nicht für alle $T > t_2$ existiert, daß also $u(x)$ entgegen der Annahme keine ganze Lösung von (1) ist.

Das folgende Beispiel soll darlegen, daß im Falle $n > 2$ die Behauptung des Satzes ohne die Voraussetzung $f(s) \geq \delta > 0$ falsch ist. Es sei

$$u(x_1, x_2, x_3) = u(r) = \begin{cases} -1575\,r^2 + 189\,r^4 - 25\,r^6 & \text{für } 0 \leq r \leq 1 \\ -1792\,r^{\frac{1}{2}} - 144\,r^{-1} + 525 & \text{für } r > 1 \end{cases}$$

mit $r = \sqrt{x_1^2 + x_2^2 + x_3^2}$. Diese Funktion ist zweimal stetig differenzierbar, und es ist

$$\Delta u = u'' + 2u'/r = \begin{cases} -9450 + 3780\,r^2 - 1050\,r^4 & \text{für } 0 \leq r \leq 1 \\ -6720\,r^{-\frac{1}{2}} & \text{für } r > 1. \end{cases}$$

Man rechnet leicht nach, daß Δu zweimal stetig differenzierbar, also u viermal stetig differenzierbar ist und daß

$$\Delta^2 u = \begin{cases} 22680 - 21000 r^2 & \text{für } 0 \leq r \leq 1 \\ 1680 r^{-\frac{1}{2}} & \text{für } r > 1 \end{cases}$$

für alle r positiv ist. Wegen $u' < 0$ hat $u(r)$ eine in $-\infty < u \leq 0$ erklärte eindeutige und stetige Umkehrfunktion $r(u)$ mit $r(0) = 0$, $r(-1411) = 1$. Hiernach ist $u(x)$ eine im ganzen R_3 viermal stetig differenzierbare Lösung der Differentialgleichung

$$\Delta^2 u = f(u) = \begin{cases} 22680 - 21000 r(u)^2 & \text{für } -1411 \leq u \leq 0 \\ 1680 r(u)^{-\frac{1}{2}} & \text{für } -\infty < u < -1411, \end{cases}$$

wobei wir für positive u über $f(u)$ noch frei verfügen können.

Damit ist der Satz vollständig bewiesen.

3. Der Satz von Nr. 2 bleibt richtig, wenn wir statt (1) die Differentialungleichung

(1a) $$\Delta^p u \geq f(u)$$

oder, was auf dasselbe hinausläuft, die Differentialgleichung

(1b) $$\Delta^p u = f(u) + F(x, u, u_{x_\nu}, \ldots)$$

mit $F \geq 0$ betrachten (über f mögen die früheren Voraussetzungen gelten). Beim Beweis ist in der Formelzeile zwischen den Formeln (15) und (16) das Gleichheitszeichen durch $\leq$ zu ersetzen. Alle übrigen Schlüsse gelten unverändert auch für diesen Fall.

Wir schließen mit einer Bemerkung über die Gleichung

$$\Delta^p u = e^u.$$

Nach unserem Satz gibt es für $n = 2$ keine ganze Lösung, während diese Frage für $n > 2$, $p \geq 2$ noch nicht geklärt ist (im Gegensatz zu $p = 1$, wo man weiß, daß es keine ganzen Lösungen gibt; s. [3]). Betrachten wir hierzu ein Beispiel! Die Funktion

$$u(x) = u(r) = a + b r^2 + \sum_{i=0}^{\infty} \frac{(-1)^i r^{2i+4}}{i!\,(2i+2)\,(2i+1)^2\,(2i+6)} \qquad (r^2 = x_1^2 + x_2^2 + x_3^2 + x_4^2)$$

ist für $x \in R_4$ analytisch, und aus $\Delta r^{2k} = 2k(2k+2) r^{2k-2}$ folgt sofort

$$\Delta^2 u = e^{-r^2} > 0.$$

Setzen wir $v(t) = u(\sqrt{t})$, so ist

$$v''(t) = \sum_{i=0}^{\infty} \frac{(-1)^i (i+1) t^i}{16 (i+3)!} = \frac{1}{16}\left[\frac{1}{6} - \frac{t}{12} + \frac{t^2}{40} - \frac{t^3}{180} + - \cdots\right]$$
$$= t^{-3}[2e^{-t} + t e^{-t} + t - 2].$$

Also gelten die Abschätzungen

$$\int_0^t v''(s)\,ds \leq \int_0^1 \frac{ds}{96} + \frac{1}{16}\int_1^\infty (3e^{-s} + s^{-2})\,ds = \frac{1}{16}\left(\frac{1}{6} + \frac{3}{e} + 1\right) < \frac{1}{6}$$

und

$$v(t) \leq a + bt + \frac{t}{6}.$$

Wählen wir $-a \geq 0$, $-b \geq 7/6$, so haben wir $v(t) \leq -t$ und

$$\Delta^2 u = e^{-r^2} = e^{-t} \geq e^{v(t)} = e^u;$$

$u(x)$ ist also eine ganze analytische Funktion, die der Ungleichung

$$\Delta^2 u \geq e^u \qquad (n = 4)$$

genügt. Dieses Beispiel legt die Vermutung nahe, daß tatsächlich ganze Lösungen von $\Delta^p u = e^u$ existieren, sobald $n > 2$, $p \geq 2$ ist.

Literatur

[1] Gröbner, W., u. N. Hofreiter: Integraltafeln II. Wien-Innsbruck 1950. — [2] Haviland, E. K.: A note un unrestricted solutions of the differential equation $\Delta u = f(u)$: J. London Math. Soc. **26**, 210—214 (1951). — [3] Walter, W.: Über ganze Lösungen der Differentialgleichung $\Delta u = f(u)$. Jahresber. DMV **57**, 94—102 (1955). — [4] Walter, W.: Mittelwertsätze und ihre Verwendung zur Lösung von Randwertaufgaben. Jahresber. DMV **59**, 93—131 (1957). — [5] Wittich, H.: Ganze Lösungen der Differentialgleichung $\Delta u = e^u$. Math. Z. **49**, 579—582 (1943/44).

Mathematisches Institut der Universität Tübingen

(Eingegangen am 16. Oktober 1956)

WALTER, W.
Math. Zeitschr. Bd. 67, S. 38—48 (1957)

Verallgemeinerte Laplace-Operatoren und Potentiale

Von
WOLFGANG WALTER

1. Verallgemeinerte Laplace-Operatoren

Es sei G ein Gebiet des n-dimensionalen Euklidischen Raumes E_n ($n \geq 2$) und $u(x) = u(x_1, \ldots, x_n)$ eine in G erklärte (reelle) Funktion. Der Oberflächenmittelwert bzw. räumliche Mittelwert von u bezüglich einer (in G gelegenen) Kugel vom Radius R und Mittelpunkt x ist definiert durch

$$m(R, x, u) = \frac{1}{\omega R^{n-1}} \oint_{|y-x|=R} u(y)\, dy \qquad (\omega = 2\pi^{n/2}/\Gamma(n/2)) \tag{1}$$

bzw.

$$M(R, x, u) = \frac{1}{\Omega R^n} \int_{|y-x|<R} u(y)\, dy = \frac{n}{R^n} \int_0^R m(r, x, u)\, r^{n-1}\, dr \qquad (n\Omega = \omega) \tag{2}$$

(falls die Integrale existieren). Mit jedem dieser Mittelwerte kann man einen verallgemeinerten Laplace-Operator bilden, und zwar ist

$$\Delta_{(1)} u(x) = 2n \lim_{R \to 0} \frac{m(R, x, u) - u(x)}{R^2} \tag{3}$$

(BLASCHKE [1]) und

$$\Delta_{(2)} u(x) = 2(n+2) \lim_{R \to 0} \frac{M(R, x, u) - u(x)}{R^2} \tag{4}$$

(PRIVALOFF [7][1]). Ist u aus der Klasse C^2[2]), so ist $\Delta_{(1)} u = \Delta_{(2)} u = \Delta u$.

In ähnlicher Weise können auch die Laplace-Operatoren höherer Ordnung verallgemeinert werden. Die (für $u \in C^{2p}$ gültige) Mittelwertentwicklung (s. [5, S. 7] [13, § 1])

$$m(R, x, u) = \sum_{i=0}^{p-1} a_{n\,i} R^{2i} \Delta^i u(x) + \int_0^R m(r, x, \Delta^p u)\, r^{n-1} H_p(r, R)\, dr \tag{5}$$

mit

$$a_{n\,i} = \frac{\Gamma(m)}{4^i\, i!\, \Gamma(m+i)} \qquad (m = n/2) \tag{6}$$

---

[1]) Zahlen in eckigen Klammern beziehen sich auf das Literaturverzeichnis am Ende der Arbeit.

[2]) Die Funktion u ist aus der Klasse C^k oder genauer aus $C^k(G)$ (kurz $u \in C^k$), wenn u und alle partiellen Ableitungen von u bis zur k-ten Ordnung in G existieren und stetig sind; u ist aus $L(G)$, wenn u über G L-integrierbar ist. Insbesondere bedeutet $h(t) \in L(0, 1)$, daß $h(t)$ über das Intervall $(0, 1)$ integrierbar ist.

und

$$H_p(r,R)=4^{1-p}(p-1)!^{-2}\int_r^R[(R^2-s^2)(s^2-r^2)]^{p-1}s^{3-n-2p}\,ds \tag{7}$$

legt die folgende Definition nahe.

Ist $u\in L(G)$ und gibt es Konstanten $\alpha_0,\ldots,\alpha_p$, so daß $m(R,x,u)$ für alle hinreichend kleinen R existiert[3]) und

$$m(R,x,u)=\sum_{i=0}^{p}a_{n\,i}R^{2i}\alpha_i+o(R^{2p})\qquad(\text{für } R\to 0) \tag{8}$$

ist, so wird

$$\Delta_{(1)}^{p}u(x)=\alpha_p$$

gesetzt (SHAPIRO [12][4]).

Das Entsprechende für den Mittelwert $M(R,x,u)$ lautet: Wenn $u\in L(G)$ ist und wenn es Konstanten $\beta_0,\ldots,\beta_p$ gibt, so daß

$$M(R,x,u)=\sum_{i=0}^{p}\frac{n}{n+2i}a_{n\,i}R^{2i}\beta_i+o(R^{2p}) \tag{9}$$

ist, wird

$$\Delta_{(2)}^{p}u(x)=\beta_p$$

gesetzt.

Ist u aus der Klasse C^{2p}, so stimmen die Operatoren $\Delta_{(1)}^{p}$, $\Delta_{(2)}^{p}$ und Δ^{p} überein. Ferner ist leicht zu sehen, daß, wenn (8) gilt, auch (9) mit $\alpha_i=\beta_i$ gültig ist. Die Definition (9) ist also allgemeiner.

Ist $p=1$ und x ein Stetigkeitspunkt von u, so ist $\alpha_0=\beta_0=u(x)$, und (8) bzw. (9) wird mit (3) bzw. (4) identisch. Ist dagegen u nur L-integrierbar, so kann in einem Punkte sehr wohl (9) für $p=1$ bestehen, ohne daß $\Delta_{(2)}u$ im Sinne von (4) existiert [dasselbe gilt für (8) und (3)]. Da aber fast überall $\beta_0=u(x)$ ist (SAKS [9, S. 118]), können wir sagen: Existiert $\Delta_{(2)}^{1}u(x)$ [gemäß (9)] in einer Punktmenge $A\subset G$, so existiert $\Delta_{(2)}u(x)$ [gemäß (4)] für $x\in A-N$, $|N|=0$[5]) [entsprechendes gilt für (8) und (3)].

Eine andere Möglichkeit zur Verallgemeinerung von Laplace-Operatoren bieten die allgemeinen Mittelwerte bezüglich einer Gewichtsfunktion $g(t)$ (s. [13, § 2]). Ist $g(t)$ für $0\le t\le 1$ erklärt, $t^{n-1}g(t)\in L(0,1)$ und $n\int_0^1 g(t)\,t^{n-1}\,dt=1$, so ist der Mittelwert $M(R,x,u,g)$ definiert durch

$$\left\{\begin{aligned}M(R,x,u,g)&=\frac{1}{\Omega R^n}\int_{|y-x|<R}u(y)\,g(|y-x|/R)\,dy\\&=\frac{n}{R^n}\int_0^R m(r,x,u)\,r^{n-1}g(r/R)\,dr.\end{aligned}\right. \tag{10}$$

[3]) Man könnte hier mit der Voraussetzung, daß u lediglich L-integrierbar ist, auskommen. Dann existiert der Oberflächenmittelwert für fast alle R eines Intervalles $(0,\delta)$ (s. [13], § 1), und man verlangt (8) nur für diejenigen R, für welche $m(R,x,u)$ existiert.

[4]) Bei SHAPIRO ist $n=2$.

[5]) $|N|$ bezeichnet das L-Maß von N.

Setzt man in das letzte Integral die Entwicklung (5) ein, so erhält man, falls $u \in C^{2p}$ ist,

$$(11)\quad \begin{cases} M(R,x,u,g) = \sum_{i=0}^{p-1} b_{ni} R^{2i} \Delta^i u(x) + \\ \qquad + \frac{n}{R^n} \int_0^R m(r,x,\Delta^p u)\, r^{n-1} \int_r^R H_p(r,s)\, s^{n-1} g(s/R)\, ds\, dr \end{cases}$$

mit

$$(12)\qquad b_{ni} = a_{ni}\, n \int_0^1 g(t)\, t^{n+2i-1}\, dt.$$

Nun sei p fest und g so gewählt, daß

$$(13)\qquad n \int_0^1 g(t)\, t^{n+2i-1}\, dt = \begin{cases} 1 & \text{für } i = 0 \\ 0 & \text{für } 0 < i < p \\ A \neq 0 & \text{für } i = p \end{cases}$$

ist. Dann gilt für $u \in C^{2p}$

$$(14)\qquad \lim_{R \to 0} \frac{M(R,x,u,g) - u(x)}{A\, a_{np} R^{2p}} = \Delta^p u(x).$$

Wir beweisen zugleich noch etwas mehr, indem wir zeigen: Ist $u \in L(G)$ und $\Delta^p_{(1)} u(x) = \alpha_p$ für einen Punkt $x \in G$, so hat der Limes (14) mit α_0 an Stelle von $u(x)$ den Wert α_p.

Es ist nämlich nach (8), (10) und (13)

$$\begin{aligned} M(R,x,u,g) &= n R^{-n} \int_0^R \left\{ \sum_{i=0}^{p} a_{ni} r^{2i} \alpha_i + o(r^{2p}) \right\} r^{n-1} g(r/R)\, dr \\ &= \alpha_0 + a_{np} A R^{2p} \alpha_p + n R^{-n} \int_0^R o(r^{2p})\, r^{n-1} g(r/R)\, dr \\ &= \alpha_0 + a_{np} A R^{2p} \alpha_p + R^{-n} o(1) \int_0^R r^{n+2p-1} |g(r/R)|\, dr \\ &= \alpha_0 + a_{np} A R^{2p} \alpha_p + R^{2p} o(1) \int_0^1 |g(t)|\, t^{n+2p-1}\, dt \\ &= \alpha_0 + a_{np} A R^{2p} \alpha_p + o(R^{2p}). \end{aligned}$$

Führt man nun den Operator Δ^p_* gemäß

$$(15)\qquad \Delta^p_* u(x) = \lim_{R \to 0} \frac{M(R,x,u,g) - u(x)}{A\, a_{np} R^{2p}}$$

ein [über g mögen die früheren Voraussetzungen, also $t^{n-1} g(t) \in L(0,1)$ und (13) gelten], so läßt sich zusammenfassend sagen:

Existiert $\Delta^p_{(1)} u(x)$ [*und ist* $\alpha_0 = u(x)$], *so existiert auch* $\Delta^p_* u(x)$, *und es ist* $\Delta^p_{(1)} u(x) = \Delta^p_* u(x)$; *ist* $u \in C^{2p}$, *so ist* $\Delta^p_* u(x) = \Delta^p u(x)$.

Ohne Beweis sei erwähnt: *Läßt sich g in der Gestalt $g = h_1 - h_2$ darstellen, wobei $h_i(t)$ in jedem Intervall $0 < \delta \leq t \leq 1$ monoton wachsend und $t^{n-1} h_i(t) \in L(0, 1)$ ist $(i = 1, 2)$, so existiert $\Delta_*^p u(x)$, sobald $\Delta_{(2)}^p u(x)$ existiert (und $\beta_0 = u(x)$ ist), und beide Werte stimmen überein.*

Als Beispiele für $g(t)$ seien

$$g_1(t) = \sum_{i=0}^{p-1} c_i \log^i t \quad \text{und} \quad g_2(t) = \sum_{i=0}^{p-1} d_i (1 - t^2)^i$$

genannt; die Konstanten c_i, d_i lassen sich so bestimmen, daß (13) gilt. Der zu g_1 gehörige Operator wurde im wesentlichen von NICOLESCO [5, S. 11][6]), der zu g_2 gehörige von GHERMANESCU [2, S. 299] angegeben.

Oft wird es zweckmäßig sein, eine engere Klasse von Gewichtsfunktionen zu betrachten, etwa nur beschränkte Funktionen oder analytische Funktionen oder Polynome zuzulassen. Dagegen erscheint es uns nicht zweckmäßig, sich ein für allemal auf eine bestimmte Gewichtsfunktion, etwa ein Polynom, festzulegen[7]).

2. Grundlösungen und Potentiale

Es sei

$$\gamma(r) = \begin{cases} -\log r & \text{für } n = 2 \\ \dfrac{r^{2-n}}{n-2} & \text{für } n > 2 \end{cases}$$

die Grundlösung für die Gleichung $\Delta u = 0$ und

$$F(x) = E_n \int f(y)\, \gamma(y - x)\, dy \tag{17}$$

das Potential einer Massenbelegung von der Dichte $f(x)$ [für $\gamma(|x|)$ schreiben wir einfach $\gamma(x)$]. Unter geeigneten Voraussetzungen über f, z.B. wenn f Hölder-stetig ist und außerhalb eines beschränkten Gebietes verschwindet, gilt dann die Poisson-Gleichung

$$\Delta F(x) = -\omega f(x). \tag{18}$$

Dieser Zusammenhang zwischen Dichtefunktion und Potential gestattet eine weitgehende Verallgemeinerung, wenn man statt Δ einen der Operatoren $\Delta_{(1)}$, $\Delta_{(2)}$ benützt.

Unter einer *positiven Massenverteilung* verstehen wir eine Mengenfunktion μ, die für alle Borel-Mengen B erklärt, nichtnegativ und volladditiv ist, für die also $\mu(B) \geq 0$ und, wenn $B_1, B_2, \ldots$ eine endliche oder unendliche Folge von paarweise punktfremden Borel-Mengen ist,

$$\mu(B_1 + B_2 + \cdots) = \mu(B_1) + \mu(B_2) + \cdots$$

---

[6]) Es ist nicht genau der Operator von NICOLESCO, sondern der Operator, den man erhält, wenn man die Nicolesco-Mittelwerte $\mu_1, \ldots, \mu_p$ statt $\mu_0, \ldots, \mu_{p-1}$ benützt (vgl. [5, S. 8ff.] und die Bemerkung in [13, § 3, Beispiel 1]).

[7]) Ist z.B. $g(t)$ ein Polynom in t^2, so ist $g(|x|/R)$ ein Polynom in x; ist g eine Funktion $P(t^2) \exp[1/(t^2 - 1)]$ (P Polynom), so hat man den Vorteil, daß der Mittelwert zu C^∞ gehört („régularisée" im Sinne von L. SCHWARTZ [11, S. 22]).

ist (RADÓ [8, S. 25]). Außerdem verlangen wir, daß die Gesamtmasse $\mu(E_n)$ endlich ist und daß es eine beschränkte Menge G mit $\mu(E_n - G) = 0$ gibt[8]). Die Funktion μ definiert ein Lebesgue-Stieltjes-Integral (SAKS [9, Chap. III], und wir nennen das Integral

$$F(x) = {}^{E_n}\!\int \gamma(y-x)\, d\mu(y) \tag{19}$$

das Potential der Massenverteilung μ (RADÓ [8, S. 27]).

Der folgende bekannte Satz kann als Verallgemeinerung der Poisson-Gleichung (18) angesehen werden.

Satz. *Ist μ eine positive Massenverteilung und $F(x)$ ihr Potential, so gilt in jedem Punkt x, in dem die Dichte $D\mu(x)$ existiert, d.h. für den*

$$\lim_{R\to 0} \frac{1}{\Omega R^n} \int\limits_{|y-x|<R} d\mu(y) = D\mu(x) < \infty \tag{20}$$

ist, die Gleichung

$$\Delta_{(1)} F(x) = \Delta_{(2)} F(x) = -\omega\, D\mu(x). \tag{21}$$

Ist insbesondere μ totalstetig, also $\mu(B) = {}^{B}\!\int f(x)\, dx$ mit $f(x) = D\mu(x) \geq 0$, so stimmen die Integrale (17) und (19) überein, und es ist fast überall

$$\Delta_{(1)} F(x) = \Delta_{(2)} F(x) = -\omega f(x)$$

(SAKS [10], POTTS [6]; vgl. auch LORENTZ [4]).

3. Es ist das *Ziel der vorliegenden Arbeit*, den entsprechenden Satz für den Operator Δ_*^p zu beweisen. Zunächst einige Erklärungen. Die Funktion

$$\gamma_p(r) = \begin{cases} A_{np} r^{2p-n},\ A_{np} = \dfrac{(-1)^{p-1}\Gamma(m-p)}{2^{2p-1}\Gamma(p)\Gamma(m)} & \text{für } n \text{ ungerade und } n > 2p \\ B_{np} r^{2p-n} \log r,\ B_{np} = \dfrac{(-1)^m}{4^{p-1}\Gamma(m)\Gamma(p)\Gamma(p-m+1)} & \text{für } n = 2, 4, \ldots, 2p \end{cases} \tag{22}$$

(hier und im folgenden ist $m = n/2$ gesetzt) ist eine Grundlösung für die Differentialgleichung $\Delta^p u = 0$ (bei JOHN [3, S. 43] steht noch ein Faktor $-\omega$ im Nenner; wir haben die aus der Potentialtheorie geläufige Normierung, bei der dieser Faktor dann in der Poisson-Gleichung auftritt, beibehalten). Wieder wird $\gamma_p(x)$ für $\gamma_p(|x|)$ geschrieben.

Das „p-Potential" ($p \geq 1$) einer positiven Massenverteilung μ ist definiert durch

$$F(x) = {}^{E_n}\!\int \gamma_p(y-x)\, d\mu(y); \tag{23}$$

ist μ totalstetig und $D\mu(x) = f(x)$, so können wir statt (23) auch

$$F(x) = {}^{E_n}\!\int f(y)\, \gamma_p(y-x)\, dy \tag{24}$$

[8]) Diese letzte Voraussetzung ist entbehrlich, wenn $n > 2$, und später bei beliebigem p, wenn $n > 2p$ ist, da dann $\gamma_p(r)$ im Unendlichen verschwindet. Ist jedoch $n \leq 2p$, so wird $\gamma_p(r)$ groß für große r, und eine solche oder ähnliche Voraussetzung wird notwendig, um die Endlichkeit des Potentials zu sichern.

schreiben. Es besteht dann, wenn f etwa Hölder-stetig ist, die iterierte Poisson-Gleichung

$$\Delta^p F(x) = -\omega f(x). \tag{25}$$

Die zu beweisende Verallgemeinerung dieses Sachverhaltes lautet:

Satz. *Ist μ eine positive Massenverteilung [mit $\mu(E_n) < \infty$ und $\mu(E_n - G) = 0$ für ein geeignetes beschränktes Gebiet G[8])], so existiert ihr Potential fast überall. Für fast alle x, genauer für jeden Punkt x, in dem die Dichte $D\mu(x)$ existiert [s. (20)], ist*

$$\Delta_*^p F(x) = -\omega D\mu(x). \tag{26}$$

Dabei ist jeder Operator Δ_*^p zugelassen, für den $g(t)\, t^{n-1} \in L(0, 1)$ ist und (13) gilt.

Nach den Bemerkungen von Nr. 1 gilt dann die Gl. (26) auch für den Operator $\Delta_{(1)}^p$ von SHAPIRO, für den allgemeineren Operator $\Delta_{(2)}^p$ und für die Operatoren von NICOLESCO und GHERMANESCU.

Beweis. Es sei N die Menge der Punkte x, in denen die Dichte existiert. Ist β eine positive Zahl $< n$ und $x \in N$, so ist nach (20) für ein festes $R > 0$

$$\begin{aligned}
&\int_{|y-x|<R} |y-x|^{\beta-n}\, d\mu(y) \\
&\quad = R^{\beta-n} \int_{|y-x|<R} d\mu(y) + \int_{|y-x|<R} \int_{|y-x|}^{R} (n-\beta)\, s^{\beta-n-1}\, ds\, d\mu(y) \\
&\quad = R^{\beta-n} \int_{|y-x|<R} d\mu(y) + (n-\beta) \int_0^R s^{\beta-n-1} \int_{|y-x|<s} d\mu(y)\, ds \\
&\quad = R^{\beta-n} \int_{|y-x|<R} d\mu(y) + O(1) \int_0^R s^{\beta-1}\, ds < \infty.
\end{aligned} \tag{27}$$

Hieraus folgt nach (22), daß das p-Potential für $x \in N$ existiert und endlich ist.

Nun sei $x_0 \in N$ und $h(x)$ eine nichtnegative Funktion aus C^∞, die in einer Umgebung des Punktes x_0 konstant gleich $D\mu(x_0)$ ist und außerhalb eines beschränkten Gebietes verschwindet. Das zu h gehörige p-Potential

$$H(x) = {}^{E_n}\!\int h(y)\, \gamma_p(y - x)\, dy$$

ist in einer Umgebung von x_0 $2p$-mal stetig differenzierbar, und es ist

$$\Delta^p H(x) = \Delta_*^p H(x) = -\omega h(x) = -\omega D\mu(x_0) \qquad \text{für } x = x_0.$$

Es genügt also zu zeigen, daß $\Delta_*^p (F - H)$ im Punkte x_0 existiert und $= 0$ ist. Nehmen wir zur Vereinfachung der Schreibweise $x_0 = 0$ an, so haben wir

$$M(R, 0, F - H, g) - F(0) + H(0) = o(R^{2p}) \qquad (\text{für } R \to 0) \tag{28}$$

unter der Voraussetzung

$$\int_{|x|<R} d\mu(x) = \Omega R^n D\mu(0) + o(R^n) \tag{29}$$

zu beweisen.

Dazu benötigen wir die Formel

$$(30)\quad G(y,r)=\oint_{|x|=r}\gamma_p(y-x)\,dx=\begin{cases}\omega\sum_{i=0}^{p-1}a_{n,i}\,r^{n+2i-1}\,\Delta^i\gamma_p(y) & \text{für } |y|>r\\ \omega\sum_{i=0}^{p-1}a_{n,i}\,r^{n-1}|y|^{2i}\,\Delta^i\gamma_p(r) & \text{für } |y|<r.\end{cases}$$

Für festes r ist $G(y,r)$ nur von $|y|$ abhängig und für $|y|\neq r$ eine p-harmonische Funktion (d.h. eine Lösung der Gleichung $\Delta^p u=0$). Die erste Zeile von (30) ist nichts anderes als die Mittelwertentwicklung (5) für $\gamma_p(y-x)$ (als Funktion von x bei festem y; man beachte, daß γ_p in der Kugel $|x|\leq r$ p-harmonisch ist). Die zweite Zeile ergibt sich aus der Tatsache, daß $G(y,r)$ für $|y|<r$ als eine nur von $|y|$ abhängende p-harmonische Funktion ein Polynom $a+b|y|^2+\cdots+c|y|^{2p-2}$ ist, das dann notwendig die angegebene Gestalt hat (man wende auf beide Seiten der Gleichung den Operator Δ^i an und setze dann $y=0$, wobei zu beachten ist, daß $\Delta^i a_{n,i}|y|^{2i}=1$ ist).

Aus den beiden Formeln

$$(31)\quad \Delta r^k=k(n+k-2)\,r^{k-2},\quad \Delta r^k\log r=[k(k+n-2)\log r+n+2k-2]\,r^{k-2}$$

folgt nach (22) leicht, daß

$$(32)\qquad \Delta^i\gamma_p(r)=C\,r^{2p-n-2i}+C_1 r^{2p-n-2i}\log r$$

mit geeigneten Konstanten C, C_1 ist und daß hierbei die Konstante C_1 nur dann von Null verschieden ist, wenn $2p-n-2i\geq 0$ und gerade ist.

Die Funktion $\gamma_p(x-y)$ genügt als Funktion von x bei festem y mit $|y|>R$ in der Kugel $|x|\leq R$ der Gleichung $\Delta^p\gamma_p(y-x)=0$. Hiernach und nach (11) bis (13) ist

$$(33)\qquad \frac{1}{\Omega R^n}\int_{|x|<R}g(|x|/R)\,\gamma_p(y-x)\,dx=\gamma_p(y)\quad\text{für } |y|>R.$$

Ist jedoch $|y|<R$, so ergibt sich mit (30).

$$(34)\quad \begin{cases}\int_{|x|<R}g(|x|/R)\,\gamma_p(y-x)\,dx=\int_0^R g(r/R)\,G(y,r)\,dr\\ =\omega\sum_{i=0}^{p-1}a_{n,i}\Big\{\Delta^i\gamma_p(y)\int_0^{|y|}g(r/R)\,r^{n+2i-1}\,dr+|y|^{2i}\int_{|y|}^R g(r/R)\,r^{n-1}\,\Delta^i\gamma_p(r)\,dr\Big\}.\end{cases}$$

Wenn man in der letzten Zeile alle Funktionen durch ihre absoluten Beträge ersetzt und beachtet, daß nach (32) für kleine r, etwa für $r<1/2$, eine Abschätzung

$$|\Delta^i\gamma_p(r)|<C_2 r^{-2i}|\gamma_p(r)|\qquad\text{mit geeignetem } C_2$$

gilt, so erhält man die Abschätzung

$$(35)\quad \begin{cases}\int_{|x|<R}|g(|x|/R)\,\gamma_p(y-x)|\,dy\leq C_3|\gamma_p(y)|\int_0^{|y|}|g(r/R)|\,r^{n-1}\,dr+\\ \qquad\qquad +C_4\int_{|y|}^R|g(r/R)\,r^{n-1}\gamma_p(r)|\,dr,\end{cases}$$

welche zeigt, daß das Integral in (33) sicher für $|y|>0$ existiert. Ist $n<2p$, so ist $\gamma_p(r)$ für $r<1/2$ beschränkt, und die rechte Seite von (35) ist (bei festem R) $<C_5$. Für $n\geq 2p$ ist $|\gamma_p(r)|$ monoton fallend, das letzte Integral wird daher vergrößert, wenn wir in ihm den Faktor $|\gamma_p(r)|$ durch $|\gamma_p(y)|$ ersetzen. In allen Fällen ist also

$$(36)\qquad \int\limits_{|x|<R} |g(|x|/R)\,\gamma_p(y-x)|\,dy \leq C_5 + C_6|\gamma_p(y)| \quad \text{für festes (kleines) } R.$$

Diese Abschätzung zeigt in Verbindung mit (27), daß auch das Integral

$$(37)\qquad \begin{cases} {}^{E_n}\!\int \int\limits_{|x|<R} |g(|x|/R)\,\gamma_p(y-x)|\,dx\,d\mu(y) \\ \qquad = \int\limits_{|x|<R} |g(|x|/R)|\,{}^{E_n}\!\int |\gamma_p(y-x)|\,d\mu(y)\,dx < \infty \end{cases}$$

ist, d.h. daß $M(R,0,F,g)$ existiert. Nach Gl. (33) ist

$$F(0)-H(0) = {}^{E_n}\!\int \gamma_p(y)\,(d\mu(y)-h(y)\,dy)$$

oder, wenn wir die Abkürzung $d\bar{\mu}(y)=d\mu(y)-h(y)\,dy$ einführen,

$$F(0)-H(0) = \int\limits_{|y|<R} \gamma_p(y)\,d\bar{\mu}(y) + \int\limits_{|y|>R} \frac{1}{\Omega R^n} \int\limits_{|x|<R} g(|x|/R)\,\gamma_p(y-x)\,dx\,d\bar{\mu}(y),$$

also wegen (37)

$$\begin{aligned} M(R,0,F-H,g) &- F(0) + H(0) \\ &= \frac{1}{\Omega R^n} \int\limits_{|x|<R} g(|x|/R)\,{}^{E_n}\!\int \gamma_p(y-x)\,d\bar{\mu}(y)\,dx - F(0) + H(0) \\ &= \frac{1}{\Omega R^n}\,{}^{E_n}\!\int \int\limits_{|x|<R} g(|x|/R)\,\gamma_p(y-x)\,dx\,d\bar{\mu}(y) - F(0) + H(0) \\ &= \int\limits_{|y|<R} \left\{ \frac{1}{\Omega R^n} \int\limits_{|x|<R} g(|x|/R)\,\gamma_p(y-x)\,dx - \gamma_p(y) \right\} d\bar{\mu}(y). \qquad (38) \end{aligned}$$

Für die geschweifte Klammer ergibt sich nach (34), wenn wir zur Abkürzung $|y|=\varrho$ setzen,

$$n R^{-n} \sum_{i=0}^{p-1} a_{n,i} \left\{ \Delta^i \gamma_p(\varrho) \int\limits_0^{\varrho} g(r/R)\,r^{n+2i-1}\,dr + \varrho^{2i} \int\limits_{\varrho}^{R} g(r/R)\,r^{n-1}\,\Delta^i\gamma_p(r)\,dr \right\} - \gamma_p(\varrho)$$

oder, da nach (13)

$$n R^{-n} \int\limits_0^R g(r/R)\,r^{n+2i-1}\,dr = 1 \quad \text{für } i=0, \; =0 \quad \text{für } 0<i<p$$

ist,

$$(39)\qquad n R^{-n} \sum_{i=0}^{p-1} a_{n,i} \int\limits_{\varrho}^{R} g(r/R)\,r^{n-1}\{\varrho^{2i}\Delta^i\gamma_p(r) - r^{2i}\Delta^i\gamma_p(\varrho)\}\,dr.$$

Betrachten wir zuerst für ein festes i den nichtlogarithmischen Teil $Cr^{2p-n-2i}$ von $\Delta^i \gamma_p(r)$. Der ihm entsprechende Teil des Integrals (38) lautet, abgesehen von einem konstanten Faktor,

$$J = R^{-n} \int_{|y|<R} \int_{\varrho}^{R} g(r/R)\, r^{n-1} \{\varrho^{2i} r^{2p-n-2i} - r^{2i} \varrho^{2p-n-2i}\}\, dr\, d\bar{\mu}(y) \quad (\varrho = |y|)$$

$$= R^{-n} \int_{|y|<R} \int_{\varrho}^{R} g(r/R)\, r^{n+2i-1} \varrho^{2i} (2p-n-4i) \int_{\varrho}^{r} s^{2p-n-4i-1}\, ds\, dr\, d\bar{\mu}(y)$$

$$(40) \qquad = (2p-n-4i)\, R^{-n} \int_{0}^{R} g(r/R)\, r^{n+2i-1} \int_{0}^{r} s^{2p-n-4i-1} \int_{|y|<s} |y|^{2i}\, d\bar{\mu}(y)\, ds\, dr.$$

Die hierbei angewandte Vertauschung der Reihenfolge der Integration ist erlaubt, da

$$\int_{0}^{R} |g(r/R)|\, r^{n+2i-1} \int_{0}^{r} s^{2p-n-4i-1} \int_{|y|<s} |y|^{2i} (d\mu(y) + h(y)\, dy)\, ds\, dr < \infty$$

ist [das über y genommene innere Integral ist nach (29) $\leq C s^{n+2i}$]. Setzen wir in (40)

$$|y|^{2i} = s^{2i} - 2i \int_{|y|}^{s} t^{2i-1}\, dt$$

ein, so spaltet sich J in zwei Integrale $J = J_1 - J_2$ auf, von denen jedes einfach abgeschätzt werden kann. Nach (29) ist

$$\int_{|y|<R} d\bar{\mu}(y) = o(R^n),$$

also

$$J_1 = (2p-n-4i)\, R^{-n} \int_{0}^{R} g(r/R)\, r^{n+2i-1} \int_{0}^{r} s^{2p-n-2i-1} \int_{|y|<s} d\bar{\mu}(y)\, ds\, dr$$

$$= R^{-n} \int_{0}^{R} g(r/R)\, r^{n+2i-1} \int_{0}^{r} s^{2p-n-2i-1}\, o(s^n)\, ds\, dr$$

$$= o(1)\, R^{-n} \int_{0}^{R} |g(r/R)|\, r^{n+2i-1} \int_{0}^{r} s^{2p-2i-1}\, ds\, dr$$

$$= o(1)\, R^{2p} \int_{0}^{1} |g(t)|\, t^{n+2p-1}\, dt = o(R^{2p})$$

und

$$J_2 = 2i(2p-n-4i)\, R^{-n} \int_{0}^{R} g(r/R)\, r^{n+2i-1} \int_{0}^{r} s^{2p-n-4i-1} \int_{|y|<s} \int_{|y|}^{s} t^{2i-1}\, dt\, d\bar{\mu}(y)\, ds\, dr$$

$$= 2i(2p-n-4i)\, R^{-n} \int_{0}^{R} g(r/R)\, r^{n+2i-1} \int_{0}^{r} s^{2p-n-4i-1} \int_{0}^{s} t^{2i-1} \int_{|y|<t} d\bar{\mu}(y)\, dt\, ds\, dr$$

$$= o(1)\, R^{-n} \int_{0}^{R} |g(r/R)|\, r^{n+2i-1} \int_{0}^{r} s^{2p-n-4i-1} \int_{0}^{s} t^{n+2i-1}\, dt\, ds\, dr = o(R^{2p}).$$

Damit ist das Integral (38), was die nichtlogarithmischen Glieder anbetrifft, in der durch (28) geforderten Weise abgeschätzt. Insbesondere ist der Satz bewiesen in all den Fällen, in denen kein logarithmisches Glied auftritt, das ist für $n \neq 2, 4, \ldots, 2p$.

Logarithmische Glieder treten nur auf, wenn $n = 2p - 2k$, $k \geq 0$ und ganz ist. Eine genauere Rechnung ergibt für diese Fälle statt (32) die Gleichung

$$\Delta^i \gamma_p(r) = B_{n,p-i} r^{2p-n-2i} \log r + C r^{2p-n-2i} \qquad (i = 0, \ldots, k); \tag{41}$$

die Konstante $B_{n,p-i}$ ist in (22) definiert, der Beweis ergibt sich leicht aus (22) und (31). Der noch abzuschätzende Rest des Integrals (38) lautet also wegen (39)

$$J' = n R^{-n} \int_{|y|<R} \sum_{i=0}^{k} a_{ni} B_{n,p-i} \int_{\varrho}^{R} g(r/R)\, r^{n-1} \{\varrho^{2i} r^{2k-2i} \log r - r^{2i} \varrho^{2k-2i} \log \varrho\}\, dr\, d\bar{\mu}(y)$$

[für $i > k$ ist $2p - 2i - n < 0$, und wir haben oben bei der Gl. (32) schon bemerkt, daß dann keine logarithmischen Glieder mehr vorhanden sind]. Diese Summe vereinfacht sich, da nach (6) und (22)

$$a_{ni} B_{n,p-i} = a_{n,k-i} B_{n,p-k+i} \qquad \text{für } n = 2p - 2k,\ 0 \leq i \leq k,\ k \text{ ganz},$$

ist, zu

$$J' = n R^{-n} \int_{|y|<R} \sum_{i=0}^{k} a_{ni} B_{n,p-i} \int_{\varrho}^{R} g(r/R)\, r^{n+2k-2i-1} \varrho^{2i} \log \frac{r}{\varrho}\, dr\, d\bar{\mu}(y) = J_0' + \cdots + J_k'.$$

Eine ähnliche Abschätzung wie oben bei den nichtlogarithmischen Gliedern ergibt

$$\begin{aligned} J_i' &= C R^{-n} \int_0^R r^{n+2k-2i-1} g(r/R) \int_{|y|<r} |y|^{2i} \log \frac{r}{|y|}\, d\bar{\mu}(y)\, dr \\ &= C R^{-n} \int_0^R r^{n+2k-2i-1} g(r/R) \int_0^r s^{-1} \int_{|y|<s} |y|^{2i} d\bar{\mu}(y)\, ds\, dr = J_i'' - J_i''' \end{aligned}$$

mit

$$\begin{aligned} J_i'' &= C R^{-n} \int_0^R r^{n+2k-2i-1} g(r/R) \int_0^r s^{2i-1} \int_{|y|<s} d\bar{\mu}(y)\, ds\, dr \\ &= o(1)\, R^{-n} \int_0^R r^{n+2k-2i-1} |g(r/R)| \int_0^r s^{2i+n-1}\, ds\, dr = o(R^{2p}) \end{aligned}$$

und

$$\begin{aligned} J_i''' &= C' R^{-n} \int_0^R r^{n+2k-2i-1} g(r/R) \int_0^r s^{-1} \int_{|y|<s} \int_{|y|}^{s} t^{2i-1}\, dt\, d\bar{\mu}(y)\, ds\, dr \\ &= C' R^{-n} \int_0^R r^{n+2k-2i-1} g(r/R) \int_0^r s^{-1} \int_0^s t^{2i-1} \int_{|y|<t} d\bar{\mu}(y)\, dt\, ds\, dr \\ &= o(1)\, R^{-n} \int_0^R r^{n+2k-2i-1} |g(r/R)| \int_0^r s^{-1} \int_0^s t^{n+2i-1}\, dt\, ds\, dr = o(R^{2p}). \end{aligned}$$

Damit ist der Satz auch in den Fällen $n = 2, 4, \ldots, 2p$ vollständig bewiesen.

Literatur

[1] BLASCHKE, W.: Ein Mittelwertsatz und eine kennzeichnende Eigenschaft des logarithmischen Potentials. Leipz. Ber. **68**, 3—7 (1916). F **46**, 742. — [2] GHERMANESCU, M.: Sur les valeurs moyennes des fonctions. Math. Ann. **119**, 288—320 (1944). — [3] JOHN, F.: Plane waves and spherical means applied to partial differential equations. New York 1955. — [4] LORENTZ, G. G.: Über die Dichte des Potentials einer räumlichen Belegung. Math. Z. **51**, 262—264 (1949). — [5] NICOLESCO, M.: Les fonctions polyharmoniques. Paris 1936. — [6] POTTS, D. H.: A note on the operators of BLASCHKE and PRIVALOFF. Bull. Amer. Math. Soc. **54**, 782—787 (1948). — [7] PRIVALOFF, I.: Sur les fonctions harmoniques. Rec. Math. Moscou **32**, 464—471 (1925). F **51**, 363. — [8] RADÓ, T.: Subharmonic functions. Berlin 1937 (Ergebnisse der Mathematik, Bd. 5, H. 1). — [9] SAKS, S.: Theory of the integral. Warschau 1937. — [10] SAKS, S.: On the operators of BLASCHKE and PRIVALOFF for subharmonic functions. Rec. Math. N. S. **9**, 451—456 (1941). R **2**, 366. — [11] SCHWARTZ, L.: Theorie des distributions I. Paris 1950. — [12] SHAPIRO, V. L.: Circular summability C of double trigonometric series. Trans. Amer. Math. Soc. **76**, 223—233 (1954). — [13] WALTER, W.: Mittelwertsätze und ihre Verwendung zur Lösung von Randwertaufgaben. Jahresber. DMV **59**, 93—131 (1957).

(F = Fortschritte der Mathematik; R = Mathematical Reviews.)

Mathematisches Institut der Universität Tübingen

(Eingegangen am 25. Oktober 1956)

BAGEMIHL, F.
Math. Zeitschr. Bd. 67, S. 49/50 (1957)

On the set of values assumed by holomorphic functions near essential singularities

By
F. BAGEMIHL

Let the point ζ in the complex plane be an isolated essential singularity of a holomorphic function $f(z)$. Then, according to WEIERSTRASS, in every neighborhood of ζ, $f(z)$ comes arbitrarily close to every complex value. It is not difficult to derive[1]) from this the fact that there exists an everywhere dense set of complex numbers, each of which is assumed by $f(z)$ in every neighborhood of ζ. PICARD's theorem, of course, overshadows this result, but, as we shall show, the idea behind its proof can be used to obtain information in a similar situation where the analogue of PICARD's theorem does not hold.

BESICOVITCH has proved[2]) that if the set E of essential singularities of a holomorphic function $f(z)$ is of linear measure zero, then, in every neighborhood of each point of E, $f(z)$ comes arbitrarily close to every complex number. SEIDEL has shown[3]), however, that there exists a holomorphic function whose essential singularities form a set of linear measure zero, and that omits a nonenumerable set of values of logarithmic measure zero[4]). We shall prove the following

Theorem. *Let the set E of essential singularities of a holomorphic function $f(z)$ be of linear measure zero. Then there exists a residual set R of complex numbers, each of which is assumed by $f(z)$ in every neighborhood of each point of E*[5]).

Proof. We shall first show that to every $\zeta \in E$ there corresponds a residual set R_ζ of complex numbers, each of which is assumed by $f(z)$ in every neighborhood of ζ.

---

[1]) *Cf.* OSGOOD, W. F.: Lehrbuch der Funktionentheorie, Bd. I, 5. Aufl., p. 329. Leipzig u. Berlin 1928. — BIEBERBACH, L.: Lehrbuch der Funktionentheorie, Bd. I, 4. Aufl., p. 150. Leipzig 1934.

[2]) BESICOVITCH, A. S.: On sufficient conditions for a function to be analytic. Proc. London Math. Soc. (2), 1—9 (1931).

[3]) SEIDEL, W.: On the distribution of values of bounded analytic functions. Trans. Amer. Math. Soc. **36**, 201—226 (1934), p. 218.

[4]) Prof. SEIDEL informs me that by making use of a result due to FROSTMAN, O.: Potentiel d'équilibre et capacité des ensembles avec quelques applications à la théorie des fonctions. Meddel. Lunds Univ. Mat. Sem. **3**, 1—118 (1935), p. 113, "logarithmic measure zero" can be replaced by "logarithmic capacity zero and infinite logarithmic measure".

[5]) Hence, the set of values that $f(z)$ omits is of first category.

By a rational disk we mean an open disk whose radius is a rational number and whose center is a complex number with its real and imaginary parts both rational. The set of all rational disks is enumerable; let its elements be denoted by $D_1, D_2, \ldots, D_n, \ldots$.

Suppose that $\zeta \in E$. We shall obtain the desired residual set R_ζ as an everywhere-dense G_δ, namely, the intersection of sets $B_1, B_2, \ldots, B_m, \ldots$, each of which is open and everywhere dense. The set B_m is constructed as follows. For every natural number n, there exists, according to BESICOVITCH's theorem, a point $d_n^{(m)} \in D_n$, and a point $z_n^{(m)}$ satisfying $0 < |z_n^{(m)} - \zeta| < 1/2m$, such that $f(z_n^{(m)}) = d_n^{(m)}$, and hence there exists an open disk $K_n^m \subset D_n$, with $d_n^{(m)} \in K_n^m$, such that every value in K_n^m is assumed by $f(z)$ in the disk C_n^m defined by $|z - z_n^{(m)}| < 1/2m$. Set $B_m = \bigcup_{n=1}^{\infty} K_n^m$. It is obvious that B_m is open, and it is also everywhere dense, because if N is any neighborhood, there exists a k such that $D_k \subset N$, and since $K_k^m \subset D_k$, $B_m \cap N \neq 0$.

Let $R_\zeta = \bigcap_{m=1}^{\infty} B_m$, and suppose that $w_0 \in R_\zeta$. Then $w_0 \in B_m$ $(m = 1, 2, 3, \ldots)$, and hence there exist natural numbers $n_1, n_2, \ldots, n_m, \ldots$ such that $w_0 \in K_{n_m}^m$ $(m = 1, 2, 3, \ldots)$. Consequently, the value w_0 is assumed by $f(z)$ in each of the disks $C_{n_m}^m$ $(m = 1, 2, 3, \ldots)$. If $r > 0$ is arbitrary, and μ is a natural number such that $1/\mu < r$, then $C_{n_\mu}^\mu$ is a subset of the neighborhood $|z - \zeta| < r$. This means that each element of R_ζ is assumed by $f(z)$ in every neighborhood of ζ.

Now let T be a countable subset of E that is everywhere dense in E. The set $R = \bigcap_{\zeta \in T} R_\zeta$ is evidently a residual set that satisfies the conclusion of the theorem.

Dept. of Math., University of Notre Dame, Notre Dame, Ind. (U.S.A.)

(Eingegangen am 11. September 1956)

AGNEW, R. P.
Math. Zeitschr. Bd. 67, S. 51—62 (1957)

Borel transforms of Tauberian series

By

R. P. AGNEW

1. Introduction. Let

$$S_n = \sum_{k=0}^{n} u_k, \qquad B(t) = e^{-t} \sum_{k=0}^{\infty} \frac{t^k}{k!} S_k \tag{1.1}$$

denote the sequence of partial sums and the Borel exponential transform of a series $\sum u_n$ of complex terms satisfying the Tauberian condition

$$\limsup_{n \to \infty} |n^{\frac{1}{2}} u_n| = L < \infty. \tag{1.2}$$

It is pur purpose to continue studies of KARAMATA [1938] and RAJAGOPAL [1953] on the behavior, for large values of n and t, of $|B(t) - S_n|$.

Let $n = n(\alpha)$ and $t = t(\alpha)$ be positive functions of a parameter α, defined for $\alpha > 0$ and such that $n(\alpha) \to \infty$ and $t(\alpha) \to \infty$ as $\alpha \to \infty$, the function $n(\alpha)$ being an integer for each $\alpha > 0$. There is then a least constant A, depending upon the functions $n(\alpha)$ and $t(\alpha)$ and which is either finite or $+\infty$, such that

$$\limsup_{\alpha \to \infty} |B(t) - S_n| \leq A\, L \tag{1.3}$$

whenever $\sum u_n$ is a series for which (1.2) holds. Our main problem is to determine this least constant A in terms of the functions $n(\alpha)$ and $t(\alpha)$, and to characterize those pairs of functions $n(\alpha)$ and $t(\alpha)$ for which A assumes its minimum value. As we shall show in section 3, if

$$\limsup_{\alpha \to \infty} |n - t|/t^{\frac{1}{2}} = M < \infty \tag{1.31}$$

then A is finite and

$$A = \left(\frac{2}{\pi}\right)^{\frac{1}{2}} \left[e^{-M^2/2} + M \int_0^M e^{-x^2/2}\, dx \right] \tag{1.32}$$

and if the left member of (1.31) is $+\infty$, then $A = +\infty$. The formula (1.32) is equivalent to

$$A = \frac{1}{(2\pi)^{\frac{1}{2}}} \int_{-\infty}^{\infty} e^{-x^2/2} |x - M|\, dx. \tag{1.33}$$

The results which we obtain imply the following two theorems.

THEOREM 1.4. *A sequence $t_0, t_1, \ldots$ for which $t_n \to \infty$ is such that*

$$\limsup_{n \to \infty} |B(t_n) - S_n| \leq (2/\pi)^{\frac{1}{2}} \limsup_{n \to \infty} |n^{\frac{1}{2}} u_n|, \tag{1.41}$$

whenever $\sum u_n$ *is a series for which the right member of* (1.41) *is finite, if and only if*

$$\lim_{n\to\infty} (t_n - n)/n^{\frac{1}{2}} = 0. \tag{1.42}$$

There is no sequence $t_0, t_1, \ldots$ *such that* $t_n \to \infty$ *and*

$$\limsup_{n\to\infty} |B(t_n) - S_n| < (2/\pi)^{\frac{1}{2}} \limsup_{n\to\infty} |n^{\frac{1}{2}} u_n| \tag{1.43}$$

whenever $\sum u_n$ *is a series for which the right member of* (1.43) *is finite and positive.*

THEOREM 1.5. *A function* $n(t)$ *such that* $n(t)$ *is an integer for each* $t>0$ *and* $n(t)\to\infty$ *as* $t\to\infty$ *is such that*

$$\limsup_{t\to\infty} |B(t) - S_{n(t)}| \leq (2/\pi)^{\frac{1}{2}} \limsup_{n\to\infty} |n^{\frac{1}{2}} u_n|, \tag{1.51}$$

whenever $\sum u_n$ *is a series for which the right member of* (1.51) *is finite, if and only if*

$$\lim_{t\to\infty} \{n(t) - t\}/t^{\frac{1}{2}} = 0. \tag{1.52}$$

There is no function $n(t)$ *such that* $n(t)\to\infty$ *and*

$$\limsup_{t\to\infty} |B(t) - S_{n(t)}| < (2/\pi)^{\frac{1}{2}} \limsup_{n\to\infty} |n^{\frac{1}{2}} u_n| \tag{1.53}$$

whenever $\sum u_n$ *is a series for which the right member of* (1.53) *is finite and positive.*

Theorems 1.4 and 1.5 imply the classic Tauberian theorem to the effect that if $B(t)\to S$ and $n^{\frac{1}{2}} u_n = o(1)$, then $S_n \to S$. theorems 1.4 and 1.5 neither imply nor are implied by the deeper Tauberian theorem to the effect that if $B(t)\to S$ and $u^{\frac{1}{2}} u_n = O(1)$, then $S_n \to S$. Theorems 1.4 and 1.5 do, however, supplement the latter theorem by showing that if $n^{\frac{1}{2}} u_n = O(1)$ and $\sum u_n$ is a series for which $\limsup |B(t)| = \infty$ [or $\limsup |S_n| = \infty$], then $\limsup |S_n| = \infty$ [or $\limsup |B(t)| = \infty$].

In section 4 to which we refer for details, it is shown that precisely the same results are obtained when (1.2) is replaced by a more general Tauberian condition of Schmidt type. The connection between our work and that of KARAMATA and RAJAGOPAL is pointed out in section 5.

2. Preliminary formulas. When (1.2) holds, all of the series below are absolutely convergent and we obtain

$$B(t) = e^{-t} \sum_{j=0}^{\infty} \frac{t^j}{j!} \sum_{k=0}^{j} u_j = \sum_{k=0}^{\infty} \left[e^{-t} \sum_{j=k}^{\infty} \frac{t^j}{j!} \right] u_k \tag{2.1}$$

and hence

$$B(t) - S_n = \sum_{k=1}^{n} - \left[1 - e^{-t} \sum_{j=k}^{\infty} \frac{t^j}{j!} \right] u_k + \sum_{k=n+1}^{\infty} \left[e^{-t} \sum_{j=k}^{\infty} \frac{t^j}{j!} \right] u_k. \tag{2.2}$$

Our next step is to throw (2.2) into a form permitting application of a general theorem on linear transformations. Let

$$x_n = n^{\frac{1}{2}} u_n \qquad n = 1, 2, 3, \ldots \tag{2.3}$$

so that the hypothesis (1.2) becomes $\limsup |x_n| = L < \infty$. For each $\alpha > 0$, let

$$a_k(\alpha) = -k^{-\frac{1}{2}} e^{-t} \sum_{j=0}^{k-1} \frac{t^j}{j!} \qquad 1 \leq k \leq n \tag{2.4}$$

and

$$a_k(\alpha) = k^{-\frac{1}{2}} e^{-t} \sum_{j=k}^{\infty} \frac{t^j}{j!} \qquad k > n. \tag{2.41}$$

Then (2.2) takes the standard form

$$B(t) - S_n = \sum_{k=1}^{\infty} a_k(\alpha)\, x_k. \tag{2.5}$$

Let, where A may be finite or $+\infty$,

$$A = \limsup_{\alpha \to \infty} \sum_{k=1}^{\infty} |a_k(\alpha)|. \tag{2.6}$$

Since

$$\lim_{\alpha \to \infty} |a_k(\alpha)| = 0 \qquad k = 1, 2, 3, \ldots \tag{2.7}$$

it follows from a lemma of AGNEW [1949] that

$$\limsup_{\alpha \to \infty} |B(t) - S_n| \leq A\, L \tag{2.8}$$

whenever (1.2) holds, and, moreover, A is the best constant in the sense that there is a series $\sum u_n$ of real terms for which (1.2) holds with $0 < L < \infty$ and equality holds in (2.8). Thus we can solve our problem by evaluating the constant A in (2.6).

3. Determination of A. From (2.4) and (2.41) we obtain, after changing orders of summation and interchanging j and k,

$$\sum_{k=1}^{\infty} |a_k(\alpha)| = e^{-t} \sum_{k=0}^{n-1} \left[\sum_{j=k+1}^{n} j^{-\frac{1}{2}} \right] \frac{t^k}{k!} + e \sum_{k=n+1}^{\infty} \left[\sum_{j=n+1}^{k} j^{-\frac{1}{2}} \right] \frac{t^k}{k!}. \tag{3.1}$$

When $0 \leq p < q$,

$$\sum_{j=p+1}^{q} j^{-\frac{1}{2}} = \int_{p}^{q} x^{-\frac{1}{2}}\, dx - \varepsilon_{p,q} = 2(q^{\frac{1}{2}} - p^{\frac{1}{2}}) - \varepsilon_{p,q} \tag{3.11}$$

where $0 < \varepsilon_{p,q} < p^{-\frac{1}{2}}$ when $p > 0$ and $0 < \varepsilon_{p,q} < 2$ when $p = 0$. From regularity of the Borel transformation, it follows that

$$\sum_{k=1}^{\infty} |a_k(\alpha)| = o(1) + 2e^{-t} \sum_{k=0}^{n-1} (n^{\frac{1}{2}} - k^{\frac{1}{2}}) \frac{t^k}{k!} + 2e^{-t} \sum_{k=n+1}^{\infty} (k^{\frac{1}{2}} - n^{\frac{1}{2}}) \frac{t^k}{k!}. \tag{3.12}$$

Therefore

$$\sum_{k=1}^{\infty} |a_k(\alpha)| = o(1) + F(\alpha) \tag{3.2}$$

where

$$F(\alpha) = 2e^{-t} \sum_{k=0}^{\infty} |k^{\frac{1}{2}} - n^{\frac{1}{2}}| \frac{t^k}{k!}. \tag{3.21}$$

As we shall show, and as can be expected from the theory of the Poisson distribution in the theory of probability, the behavior of $F(\alpha)$ depends upon the behavior of the function $\omega = \omega(\alpha)$ defined by

$$\omega = (n - t)/t^{\frac{1}{2}} \tag{3.3}$$

or

$$n = t + \omega t^{\frac{1}{2}}. \tag{3.31}$$

We consider separately the cases in which $\limsup |\omega(\alpha)|$ is infinite and finite.

THEOREM 3.4. *If*

$$\limsup_{\alpha \to \infty} |\omega(\alpha)| = \infty \tag{3.41}$$

then $A = \limsup F(\alpha) = \infty$.

To prove this, let a be a constant for which $a > 1$. From (3.21) and (3.31) we obtain, when α is sufficiently large,

$$\begin{cases} F(\alpha) > 2e^{-t} \sum\limits_{|k-t| < at^{\frac{1}{2}}} |k^{\frac{1}{2}} - n^{\frac{1}{2}}| \frac{t^k}{k!} = 2e^{-t} \sum\limits_{|k-t| < at^{\frac{1}{2}}} \frac{|k-n|}{k^{\frac{1}{2}} + n^{\frac{1}{2}}} \frac{t^k}{k!} > \\ > \frac{2}{(t - at^{\frac{1}{2}})^{\frac{1}{2}}} e^{-t} \sum\limits_{|k-t| < at^{\frac{1}{2}}} |k - (t + \omega t^{\frac{1}{2}})| \frac{t^k}{k!} \end{cases} \tag{3.42}$$

and hence

$$F(\alpha) > 2 \frac{(|\omega| - a)\, t^{\frac{1}{2}}}{(t - at^{\frac{1}{2}})^{\frac{1}{2}}} e^{-t} \sum_{|k-t| < at^{\frac{1}{2}}} \frac{t^k}{k!}. \tag{3.43}$$

The elementary formulas

$$e^t = \sum_{k=0}^{\infty} \frac{t^k}{k!}, \quad t e^t = \sum_{k=0}^{\infty} \frac{k t^k}{k!}, \quad t \frac{d}{dt} t e^t = \sum_{k=0}^{\infty} \frac{k^2 t^k}{k!} \tag{3.44}$$

give, for each real x, the formula

$$e^{-t} \sum_{k=0}^{\infty} (k - x)^2 \frac{t^k}{k!} = (t - x)^2 + t. \tag{3.441}$$

Using this formula with $x = t$ gives

$$e^{-t} \sum_{|k-t| \geq at^{\frac{1}{2}}} (k - t)^2 \frac{t^k}{k!} < t \tag{3.442}$$

and

$$e^{-t} \sum_{|k-t| \geq at^{\frac{1}{2}}} a^2 t \frac{t^k}{k!} < t \tag{3.443}$$

so

$$e^{-t}\sum_{|k-t|\geq at^{\frac{1}{2}}}\frac{t^k}{k!}<\frac{1}{a^2} \tag{3.444}$$

and therefore

$$e^{-t}\sum_{|k-t|<at^{\frac{1}{2}}}\frac{t^k}{k!}>1-\frac{1}{a^2}. \tag{3.445}$$

This and (3.43) imply that

$$F(\alpha)>2\frac{(|\omega|-a)\,t^{\frac{1}{2}}}{(t-at^{\frac{1}{2}})^{\frac{1}{2}}}\quad\left(1-\frac{1}{a^2}\right) \tag{3.45}$$

and the conclusion of Theorem 3.4 follows.

THEOREM 3.5. *If*

$$\limsup_{\alpha\to\infty}|\omega(\alpha)|<\infty \tag{3.51}$$

then

$$F(\alpha)=o(1)+\left(\frac{2}{\pi}\right)^{\frac{1}{2}}\left[e^{-\omega^2/2}+|\omega|\int_0^{|\omega|}e^{-x^2/2}\,dx\right]. \tag{3.52}$$

To prove this, let $\varepsilon>0$. With ω defined by (3.31), the hypothesis (3.51) implies that $\lim n^{-1}t=1$ and that we can choose a positive constant a so great that

$$\limsup_{\alpha\to\infty}2n^{-\frac{1}{2}}(\omega^2+1)^{\frac{1}{4}}t^{\frac{1}{2}}a^{-1}<\varepsilon. \tag{3.53}$$

We now split $F(\alpha)$ into the sum of two terms of which the first is

$$F_1(\alpha)=2e^{-t}\sum_{|k-t|>at^{\frac{1}{2}}}|k^{\frac{1}{2}}-n^{\frac{1}{2}}|\,\frac{t^k}{k!}. \tag{3.54}$$

Since

$$|k^{\frac{1}{2}}-n^{\frac{1}{2}}|=\frac{|k-n|}{k^{\frac{1}{2}}+n^{\frac{1}{2}}}\leq\frac{|k-n|}{n^{\frac{1}{2}}} \tag{3.55}$$

we obtain

$$F_1(\alpha)\leq 2n^{-\frac{1}{2}}e^{-t}\sum_{|k-t|>at^{\frac{1}{2}}}\left[(k-n)^2\frac{t^k}{k!}\right]^{\frac{1}{2}}\left[\frac{t^k}{k!}\right]^{\frac{1}{2}}. \tag{3.56}$$

Use of the Schwarz inequality gives

$$|F_1(\alpha)|\leq 2n^{-\frac{1}{2}}\left[e^{-t}\sum_{|k-t|>at^{\frac{1}{2}}}(k-n)^2\frac{t^k}{k!}\right]^{\frac{1}{2}}\left[e^{-t}\sum_{|k-t|<at^{\frac{1}{2}}}\frac{t^k}{k!}\right]^{\frac{1}{2}}. \tag{3.57}$$

Using (3.31), (3.441), (3.444), and (3.53) gives

$$\limsup_{\alpha\to\infty}F_1(\alpha)<\varepsilon. \tag{3.58}$$

Using (3.21) and (3.54) gives

$$F(\alpha)=F_1(\alpha)+2e^{-t}\sum_{|k-t|\leq at^{\frac{1}{2}}}|k^{\frac{1}{2}}-n^{\frac{1}{2}}|\,\frac{t^k}{k!}. \tag{3.6}$$

For the range of values of k in the sum in (3.6),

$$2\left|k^{\frac{1}{2}}-n^{\frac{1}{2}}\right| = t^{-\frac{1}{2}}\left|k-n\right|\left\{\frac{2t^{\frac{1}{2}}}{k^{\frac{1}{2}}+n^{\frac{1}{2}}}\right\} \tag{3.61}$$

and the last factor in braces converges uniformly to 1 as $\alpha\to\infty$.

Therefore

$$F(\alpha) = o(1) + F_1(\alpha) + t^{-\frac{1}{2}}e^{-t}\sum_{|k-t|\leq at^{\frac{1}{2}}}\left|k-n\right|\frac{t^k}{k!}. \tag{3.62}$$

Putting

$$F_2(\alpha) = t^{-\frac{1}{2}}e^{-t}\sum_{|k-t|>at^{\frac{1}{2}}}\left|k-n\right|\frac{t^k}{k!} \tag{3.63}$$

and using the estimates following (3.55) gives

$$\limsup_{\alpha\to\infty} F_2(\alpha) < \varepsilon. \tag{3.64}$$

Moreover

$$F(\alpha) = o(1) + F_1(\alpha) - F_2(\alpha) + t^{-\frac{1}{2}}e^{-t}\sum_{k=0}^{\infty}\left|k-n\right|\frac{t^k}{k!} \tag{3.65}$$

and because of (3.58) and (3.64) we conclude that

$$F(\alpha) = o(1) + G(\alpha) \tag{3.66}$$

where

$$G(\alpha) = t^{-\frac{1}{2}}e^{-t}\sum_{k=0}^{\infty}\left|k-n\right|\frac{t^k}{k!}. \tag{3.7}$$

Splitting the sum in (3.7) into parts for which $0\leq k\leq n$ and $k>n$ and using (3.31) gives

$$G(\alpha) = g_1(\alpha) + g_2(\alpha) - \omega \tag{3.71}$$

where

$$g_1(\alpha) = 2t^{\frac{1}{2}}e^{-t}t^n/n!, \tag{3.72}$$

$$g_2(\alpha) = 2\omega e^{-t}\sum_{k=0}^{n}\frac{t^k}{k!}. \tag{3.73}$$

Use of (3.72), (3.31) and STIRLING's formula for $n!$ gives

$$g_1(\alpha) = o(1) + (2/\pi)^{\frac{1}{2}}e^{-\omega^2/2}. \tag{3.74}$$

The fundamental formula

$$e^{-t}\sum_{k\leq t+\omega t^{\frac{1}{2}}}\frac{t^k}{k!} = o(1) + (2\pi)^{-\frac{1}{2}}\int_{-\infty}^{\omega}e^{-x^2/2}\,dx \tag{3.75}$$

gives

$$g_2(\alpha) = o(1) + \omega\left(\frac{2}{\pi}\right)^{\frac{1}{2}}\int_{-\infty}^{\omega}e^{-x^2/2}\,dx \tag{3.76}$$

and hence

$$g_2(\alpha) - \omega = o(1) + |\omega| \left(\frac{2}{\pi}\right)^{\frac{1}{2}} \int_0^{|\omega|} e^{-x^2/2}\, dx. \tag{3.77}$$

From (3.66), (3.71), (3.74), and (3.76) we obtain (3.52) and Theorem 3.5 is proved.

Because the result may be of interest in other connections, we pause to give the following Theorem.

THEOREM 3.8. *If n and t become infinite in such a way that the quantity ω defined by*

$$n = t + \omega t^{\frac{1}{2}} \tag{3.81}$$

is bounded, then

$$2e^{-t} \sum_{k=0}^{\infty} |k^{\frac{1}{2}} - n^{\frac{1}{2}}| \frac{t^k}{k!} = o(1) + \frac{1}{(2\pi)^{\frac{1}{2}}} \int_{-\infty}^{\infty} |x - \omega|\, e^{-x^2/2}\, dx \tag{3.82}$$

and

$$t^{-\frac{1}{2}} e^{-t} \sum_{k=0}^{\infty} |k - n| \frac{t^k}{k!} = o(1) + \frac{1}{(2\pi)^{\frac{1}{2}}} \int_{-\infty}^{\infty} |x - \omega|\, e^{-x^2/2}\, dx. \tag{3.83}$$

If $\limsup |\omega| = \infty$, then the superior limits of the left members of (3.82) and (3.83) are infinite.

To prove this Theorem, we observe that the left member of (3.82) is the function $F(\alpha)$ in (3.21) and (3.52), and that for each ω the integral in (3.82) is equal to the last part of (3.52). In case ω is bounded, (3.82) therefore follows from Theorem 3.5 and (3.83) then follows from (3.66) and (3.7); in case ω is unbounded the conclusion of Theorem 3.8 follows from Theorem 3.4. In case ω has a limit say, $\lim \omega = \omega_0$, (3.82) is valid and reduces to

$$\lim 2e^{-t} \sum_{k=0}^{\infty} |k^{\frac{1}{2}} - n^{\frac{1}{2}}| \frac{t^k}{k!} = \frac{1}{(2\pi)^{\frac{1}{2}}} \int_{-\infty}^{\infty} |x - \omega_0|\, e^{-x^2/2}\, dx. \tag{3.84}$$

Since the last term in (3.52) increases with $|\omega|$, and attains its minimum value $(2/\pi)^{\frac{1}{2}}$ when and only when $\omega = 0$, Theorems 3.4 and 3.5 show precisely how the best constants A defined in the introduction depend upon the given functions $n(\alpha)$ and $t(\alpha)$. In case $\limsup |\omega(\alpha)| = \infty$, that is

$$\limsup |n - t|/t^{\frac{1}{2}} = \infty, \tag{3.9}$$

we see from Theorem 3.4 that $A = \infty$. In case

$$\limsup |n - t|/t^{\frac{1}{2}} = M < \infty, \tag{3.91}$$

we see from Theorem 3.5 that

$$A = \left(\frac{2}{\pi}\right)^{\frac{1}{2}} \left[e^{-M^2/2} + M \int_0^M e^{-x^2/2}\, dx\right]. \tag{3.92}$$

This constant A attains its minimum value

$$A_0 = (2/\pi)^{\frac{1}{2}} = 0 \cdot 79788\ 45608\ 02865\ 35587\ 98921 \tag{3.93}$$

when and only when $M = 0$, that is

$$\lim (n - t)/t^{\frac{1}{2}} = 0. \tag{3.94}$$

4. Schmidt conditions. It is well known and easy to prove that if a series $\sum u_n$ satisfies the Tauberian condition (1.2), then its partial sums $S_n = u_0 + u_1 + \cdots + u_n$ satisfy the more general Tauberian condition

$$\limsup_{p \to \infty} \operatorname*{maximum}_{|q-p| \leq \lambda p^{\frac{1}{2}}} |S_q - S_p| \leq L\lambda, \qquad \lambda > 0, \tag{4.1}$$

of Schmidt type.

Letting $n(\alpha)$ and $t(\alpha)$ be defined as in section 1, we determine the least constant A', which depends upon the functions $n(\alpha)$ and $t(\alpha)$ and which is either finite or $+\infty$, such that

$$\limsup_{\alpha \to \infty} |B(t) - S_n| \leq A'L \tag{4.11}$$

whenever $\sum u_n$ is a series for which (4.1) holds.

Since each series satisfying (1.2) also satisfies (4.1), it is clear from the definitions of A and A' in (1.3) and (4.11) that $A' \geq A$. If for some functions $n(\alpha)$ and $t(\alpha)$ it were true that $A' > A$, it would perhaps be difficult to determine A' because the condition (4.1) is not as tractable as the condition (1.2). However, as we shall show, matters are greatly simplified by the fact that we always have $A' = A$. Since we know that $A' \geq A$, we can prove this by proving that $A' \leq A$. In case the function $\omega = \omega(\alpha)$ defined by (3.3) is such that $\limsup |\omega(\alpha)| = \infty$, we have $A = \infty$ and the inequality $A' \leq A$ is automatically satisfied. Thus we can complete the proof that $A' = A$ by proving the following special result which we formulate as a lemma.

Lemma 4.2. *If $n(\alpha)$ and $t(\alpha)$ are such that* $\limsup |\omega(t)| = M < \infty$ *and if $\sum u_n$ is a series satisfying* (4.1), *then* (4.11) *holds when A' is the right member of* (1.32) *and* (1.33).

To prove this lemma, let $H(\alpha) = B(t) - S_n$ so that

$$H(\alpha) = e^{-t} \sum_{k=0}^{\infty} \frac{t^k}{k!} (S_k - S_n). \tag{4.21}$$

Let $\varepsilon > 0$ and choose a positive constant a so great that $a > 1$ and (3.53) holds. Then

$$|H(\alpha)| \leq H_1(\alpha) + H_2(\alpha) \tag{4.22}$$

where

$$H_1(\alpha) = e^{-t} \sum_{|k-t| > at^{\frac{1}{2}}} \frac{t^k}{k!} |S_k - S_n| \tag{4.23}$$

and

$$H_2(\alpha) = e^{-t} \sum_{|k-t| \leq at^{\frac{1}{2}}} \frac{t^k}{k!} |S_k - S_n|. \tag{4.24}$$

SCHMIDT [1925, page 232] has obtained a result which implies that if a series $\sum u_n$ satisfies (4.1), then there is a constant C such that the inequality

$$|S_k - S_n| \leq C\,|k^{\frac{1}{2}} - n^{\frac{1}{2}}| \tag{4.25}$$

holds whenever k and n are nonnegative integers for which $|k-n| > n^{\frac{1}{2}}$. Since $a > 1$ and $\omega(\alpha)$ is bounded so that $t(\alpha)/n(\alpha) \to 1$ as $\alpha \to \infty$, it follows that if α is sufficiently great then the inequality $|k-t| > a t^{\frac{1}{2}}$ implies that $|k-t| > n^{\frac{1}{2}}$ and hence that

$$H_1(\alpha) \leq C\,e^{-t} \sum_{|k-t| > a t^{\frac{1}{2}}} |k^{\frac{1}{2}} - n^{\frac{1}{2}}| \, \frac{t^k}{k!}. \tag{4.26}$$

Use of the formulas (3.54) to (3.58) then shows that

$$\limsup_{\alpha \to \infty} H_1(\alpha) \leq C\,\varepsilon. \tag{4.27}$$

To estimate $H_2(\alpha)$, let N be an integer so great that $N > 1$ and

$$\frac{a}{N} \int_{-a}^{a} e^{-x^2/2}\,dx < \varepsilon. \tag{4.3}$$

Starting with (4.24) then gives

$$H_2(\alpha) = e^{-t} \sum_{p=-N}^{N-1} \sum_{k \in E(p,\alpha)} \frac{t^k}{k!}\,|S_k - S_n| \tag{4.4}$$

where $E(p, \alpha)$ is the set of nonnegative integers k for which

$$t + \frac{p}{N}\,a\,t^{\frac{1}{2}} \leq k \leq t + \frac{p+1}{N}\,a\,t^{\frac{1}{2}}. \tag{4.41}$$

From (4.4) we obtain

$$H_2(\alpha) \leq e^{-t} \sum_{p=-N}^{N-1} \left[\sum_{k \in E(p,\alpha)} \frac{t^k}{k!} \right] f_p(\alpha) \tag{4.42}$$

where

$$f_p(\alpha) = \max_{k \in E(p,\alpha)} |S_k - S_n|. \tag{4.43}$$

From (4.43) and the definitions of $\omega(\alpha)$ and $E(p, \alpha)$ in (3.3) and (4.41) we obtain

$$f_p(\alpha) = \underset{(pN^{-1}a-\omega)t^{\frac{1}{2}} \leq k-n \leq [(p+1)N^{-1}a-\omega]t^{\frac{1}{2}}}{\text{maximum}} |S_k - S_n| \tag{4.44}$$

and hence

$$f_p(\alpha) \leq \underset{|k-n| \leq x(p,\alpha)\,t^{\frac{1}{2}}}{\text{maximum}} |S_k - S_n| \tag{4.45}$$

where

$$x(p, \alpha) = \max\left\{ \left| \frac{p}{N}\,a - \omega \right|, \left| \frac{p+1}{N}\,a - \omega \right| \right\}. \tag{4.46}$$

We estimate the functions $f_p(\alpha)$ by use of the following lemma.

Lemma 4.5. *If* $\sum u_n$ *satisfies the* Schmidt *condition* (4.1) *and if*

$$f(\alpha) = \underset{|k-n(\alpha)| \leq x(\alpha)\,[t(\alpha)]^{\frac{1}{2}}}{\text{maximum}} |S_k - S_{n(\alpha)}| \tag{4.51}$$

where $n(\alpha)$, $t(\alpha)$, *and* $x(\alpha)$ *are positive functions for which* $n(\alpha) \to \infty$ *and* $t(\alpha)/n(\alpha) \to 1$ *as* $\alpha \to \infty$ *and* $x(\alpha)$ *is bounded. Then to each* $\varepsilon > 0$ *corresponds a number* α_0 *such that*

$$f(\alpha) < L\,x(\alpha) + \varepsilon, \qquad \alpha > \alpha_0. \tag{4.52}$$

To prove this lemma, suppose that the conclusion is false. Then for some $\varepsilon > 0$ there is a sequence $\alpha_1, \alpha_2, \ldots$ such that $\alpha_p \to \infty$ and

$$f(\alpha_p) \geq L\,x(\alpha_p) + \varepsilon, \qquad p = 1, 2, \ldots. \tag{4.53}$$

Since the function $x(\alpha)$ is bounded, the sequence α_p can be chosen such that $\lim_{p \to \infty} x(\alpha_p)$ exists and is a number x_0. Let $h > 0$. We then have, for each sufficiently great p,

$$\underset{|k-n(\alpha_p)| \leq (x_0+h)(1+h)[n(\alpha_p)]^{\frac{1}{2}}}{\text{maximum}} |S_k - S_{n(\alpha_p)}| \geq L(x_0 - h) + \varepsilon, \tag{4.54}$$

and use of the hypothesis (4.1) gives

$$L(x_0 + h)(1 + h) \geq L(x_0 - h) + \varepsilon. \tag{4.55}$$

Since (4.55) holds for each $h > 0$, it implies that $\varepsilon \leq 0$ and gives a contradiction that establishes Lemma 4.5.

Applying Lemma 4.5 to the functions $f_p(\alpha)$ in (4.45), we find that there is a number α_0 such that, for each integer p in the interval $-N \leq p \leq N-1$,

$$f_p(\alpha) \leq L\,x(p, \alpha) + \varepsilon, \qquad \alpha > \alpha_0. \tag{4.6}$$

This and (4.42) imply that

$$H_2(\alpha) \leq e^{-t} \sum_{p=-N}^{N-1} \left[\sum_{k \in E(p,\alpha)} \frac{t^k}{k!} \right] [L\,x(p, \alpha) + \varepsilon] \tag{4.61}$$

when $\alpha > \alpha_0$. Hence

$$H_2(\alpha) \leq \varepsilon + L \sum_{p=-N}^{N-1} \left[e^{-t} \sum_{k \in E(p,\alpha)} \frac{t^k}{k!} \right] x(p, \alpha) \tag{4.62}$$

when $\alpha > \alpha_0$. Using the definition of $E(p, \alpha)$ in (4.41) and the fundamental formula (3.75), we see that the quantity in square brackets in the right member of (4.62) converges to

$$\frac{1}{(2\pi)^{\frac{1}{2}}} \int_{pN^{-1}a}^{(p+1)N^{-1}a} e^{-u^2/2}\,du \tag{4.63}$$

as $\alpha \to \infty$. Hence we can apply the simple fact that if

$$f(\alpha) = \sum_{k=1}^{n} g_k(\alpha)\,h_k(\alpha) \tag{4.64}$$

where all of the functions involved are nonnegative and bounded and $\lim_{\alpha\to\infty} g_k(\alpha) = G_k$ for each k, then

$$\limsup_{\alpha\to\infty} f(\alpha) = \limsup_{\alpha\to\infty} \sum_{k=1}^{n} G_k h_k(\alpha). \tag{4.65}$$

This and (4.62) give

$$\limsup_{\alpha\to\infty} H_2(\alpha) \leq \varepsilon + (2\pi)^{-\frac{1}{2}} L \limsup_{\alpha\to\infty} H_3(\alpha) \tag{4.7}$$

where

$$H_3(\alpha) = \sum_{p=-N}^{N-1} \left[\int_{pN^{-1}a}^{(p+1)N^{-1}a} e^{-u^2/2}\, du \right] x(p,\alpha). \tag{4.71}$$

From (4.46) we see that

$$x(p,\alpha) = |x_p - \omega| \tag{4.72}$$

where x_p is, for each p and α, a number for which $pN^{-1}a \leq x_p \leq (p+1)N^{-1}a$. Hence when $pN^{-1}a \leq x \leq (p+1)N^{-1}a$ we have $|x - x_p| \leq N^{-1}a$ and consequently

$$x(p,\alpha) \leq N^{-1}a + |x - \omega|. \tag{4.73}$$

This and (4.71) give

$$H_3(\alpha) = \sum_{p=-N}^{N-1} \int_{pN^{-1}a}^{(p+1)N^{-1}a} e^{-x^2/2} [N^{-1}a + |x-\omega|]\, dx \tag{4.74}$$

and use of (4.3) gives

$$H_3(\alpha) \leq \varepsilon + \int_{-a}^{a} e^{-x^2/2} |x - \omega|\, dx. \tag{4.75}$$

Using (4.22), (4.27), (4.7) and (4.75) gives

$$\limsup_{\alpha\to\infty} |H(\alpha)| \leq C\varepsilon + \varepsilon + \frac{L}{(2\pi)^{\frac{1}{2}}} \limsup_{\alpha\to\infty} \varepsilon + \int_{-\infty}^{\infty} e^{-x^2/2} |x - \omega|\, dx. \tag{4.76}$$

Since (4.76) holds for each $\varepsilon > 0$, it implies that

$$\limsup_{\alpha\to\infty} H(\alpha) = \frac{L}{(2\pi)^{\frac{1}{2}}} \limsup_{\alpha\to\infty} \int_{-\infty}^{\infty} e^{-x^2/2} |x - \omega|\, dx. \tag{4.77}$$

Therefore (4.11) holds when

$$A' = \frac{1}{(2\pi)^{\frac{1}{2}}} \int_{-\infty}^{\infty} e^{-x^2/2} |x - M|\, dx \tag{4.78}$$

where M is defined as before by the formula

$$M = \limsup_{\alpha\to\infty} |\omega(t)| = \limsup_{\alpha\to\infty} |n - t|/t^{\frac{1}{2}}. \tag{4.79}$$

Since the right member of (4.78) is the right member of (1.33), this completes the proof of Lemma 4.2. The equality $A' = A$ is therefore established.

5. Conclusion. Since (3.94) obviously holds when $n = [t]$, the greatest integer in t, our results include that of RAJAGOPAL [1953] to the effect that

$$\limsup_{t\to\infty} |B(t) - S_{[t]}| \leq (2/\pi)^{\frac{1}{2}} \limsup_{n\to\infty} |n^{\frac{1}{2}} u_n| \tag{5.1}$$

whenever $\sum u_n$ is a series for which the right member is finite. RAJAGOPAL [1953] showed, by giving a specific numerical series $\sum u_n$ suuplied to him by T. VIJAYARAGHAVAN, that the constant $(2/\pi)^{\frac{1}{2}}$ is the least constant for which (5.1) is valid. Without undertaking to determine best constants, KARAMATA [1938] obtained inequalities involving $|B(t) - S_n|$ for the case in which $t = n + an^{\frac{1}{2}}$, a is a constant, and $\sum u_n$ satisfies a Tauberian condition of Schmidt type. KARAMATA obtained Tauberian Theorems which employ the hypothesis that $\lim B(t)$ exists and which neither imply nor are implied by our Theorems involving (1.3). The papers of KARAMATA and RAJAGOPAL contain references to earlier works on the subject.

References

AGNEW, R. P. [1949]: Abel transforms and partial sums of Tauberian series. Ann. of Math. (2) **50**, 110—117. — KARAMATA, JOVAN [1938]: O jednom opštem O-inversnom stavu. Rad Jugoslavenske Akademije Znanosti i Umjetnosti, Knjiga 261, Matematicko-Prirodoslovni Razred 81, pp. 1—22. — RAJAGOPAL, C. T. [1953]: A generalization of TAUBER's Theorem and some Tauberian constants. Math. Z. **57**, 405—414. — SCHMIDT, ROBERT [1925]: Die Umkehrsätze des Borelschen Summierungsverfahrens. Schriften der Königsberger Gelehrten Gesellschaft, Naturwiss. Kl. **1**, 205—256.

Cornell University, Ithaca, New York

(Eingegangen am 29. September 1956)

Laugwitz, D.
Math. Zeitschr. Bd. 67, S. 63—74 (1957)

Zur Differentialgeometrie der Hyperflächen in Vektorräumen und zur affingeometrischen Deutung der Theorie der Finsler-Räume

Von

Detlef Laugwitz

Die vorliegenden Untersuchungen sind veranlaßt von einer von E. R. Lorch bemerkten Tatsache: Mit einer glatten Hyperfläche $\mathfrak{F}$ in einem reellen Vektorraum (= radialaffinen Raum) ist eine Riemannsche Metrik in diesem Raume gegeben, die nach ihrer Definition affininvariant mit der Fläche $\mathfrak{F}$ verknüpft ist. Diese Maßbestimmung, die zunächst bei Hyperflächen im Hilbertschen Raum für funktionalanalytische Anwendungen eingeführt wurde, besitzt auch im endlichdimensionalen Fall interessante differentialgeometrische Eigenschaften, deren Untersuchung in einer früheren Arbeit begonnen wurde[1]). Hier setzen wir diese Untersuchung fort, indem wir die auf der Hyperfläche $\mathfrak{F}$ selbst induzierte Riemannsche Flächenmetrik betrachten.

Wir geben zunächst eine geometrische Definition dieser Flächenmetrik (§ 1) und zeigen dann in § 2, daß diese Metrik im endlichdimensionalen Fall (bis auf einen unwesentlichen Faktor) mit der bekannten Flächenmetrik der radialaffinen Geometrie von E. Salkowski[2]) zusammenfällt. Es wird hier also ein neuer Aufbau der radialaffinen Flächentheorie ermöglicht, dessen Grundzüge wegen der von Salkowskis Darstellung wesentlich verschiedenen Methode hier kurz skizziert werden. Außerdem ermöglicht unsere Methode eine Erweiterung auf unendlichdimensionale Räume, also auf die besonders von Lorch[3]) betrachteten Fälle; eine solche Verallgemeinerung wäre mit der Salkowskischen Methode, in der n-reihige Determinanten eine wichtige Rolle spielen, nicht möglich. Diese Erweiterung ergibt sich mittels der früher entwickelten Tensorrechnung für unendlichdimensionale Mannigfaltigkeiten[4]); dieser Kalkül — bei dem man das Erscheinungsbild der Formeln gegenüber der klassischen Tensorrechnung nicht abzuändern braucht — bringt bei dem vorliegenden Gegenstand auch für endlichdimensionale Vektorräume den

[1]) Laugwitz, D., and E. R. Lorch: Riemannian metrics associated with convex bodies in normed spaces. Amer. J. Math. **78**, 889—894 (1956).

[2]) Salkowski, E.: Affine Differentialgeometrie. Berlin-Leipzig 1934.

[3]) A curvature study of convex bodies in Banach spaces. Annali mat. pura appl. (IV) **34**, 105—112 (1953).

[4]) Math. Z. **61**, 100—118 (1954); für unendlichdimensionale Minkowskische und Finslersche Räume: Grundlagen für die Geometrie der unendlichdimensionalen Finsler-Räume. Annali mat. pura appl. (IV) **41**, 21—41 (1955).

begrifflich befriedigenderen Vorteil, daß ein basisfreies Rechnen (also eine Operation mit Vektoren selbst und nicht mit Komponenten) möglich ist. Demgemäß sind alle Formeln so gehalten, daß sie — sofern nicht ausdrücklich etwas anderes gesagt ist — sowohl im Sinne der üblichen Tensorrechnung als auch gemäß der erwähnten Verallgemeinerung[4]) für unendlichdimensionale Räume gültig sind.

Radialaffine Räume haben für die Differentialgeometrie ein besonderes Interesse, weil sie als Tangentialräume differenzierbarer Mannigfaltigkeiten auftreten. Insbesondere ist ein Finslerscher Raum[5]) dadurch definiert, daß im Tangentialraum eines jeden seiner Punkte eine Hyperfläche als Indikatrix ausgezeichnet ist, wodurch dieser Tangentialraum zu einem Minkowskischen Raum wird. Unsere Auffassung der radialaffinen Flächentheorie läßt sich unmittelbar auf die Indikatrix anwenden — der Fundamentaltensor der Riemannschen Metrik von LORCH ist formal identisch mit dem Fundamentaltensor der Minkowskischen Maßbestimmung —, und dadurch wird sich eine affingeometrische Deutung der von E. CARTAN in der Finslerschen Geometrie eingeführten Tensoren ergeben[6]). Die von O. VARGA kürzlich[7]) durchgeführte Krümmungstheorie der Indikatrix eines Minkowskischen Raumes läßt sich — wie schon in einer früheren Arbeit, l. c.[1]), angedeutet wurde — hier einbauen und erweist sich als im wesentlichen identisch mit der Krümmungstheorie der radialaffinen Flächengeometrie.

Es darf wohl behauptet werden, daß von unserem Standpunkt eine geometrisch befriedigende Synthese der beiden Auffassungen des Finslerschen Raumes — Linienelement-Auffassung von E. CARTAN einerseits und sog. lokal-Minkowskische Auffassung von H. RUND und anderen — ermöglicht wird. Danach wird man den Finslerschen Raum auffassen als einen Punktraum, in dessen Tangentialräumen vermöge der Indikatrix einmal je eine Minskowskische Metrik (Norm) definiert ist, wobei zum anderen aber durch die Raummetrik der radialaffinen Geometrie auch ein natürlicher geometrischer Einbau der Auffassung von CARTAN möglich ist. Dadurch bleibt der anschauliche Charakter eines Punktraumes gewahrt, es wird aber der wegen seiner Linearisierungen bequeme Cartansche Kalkül mit übernommen.

Es sei noch darauf hingewiesen, daß hier — anders als in SALKOWSKIs Affingeometrie — eine Metrik nicht nur in der Fläche, sondern im ganzen Raume definiert ist, und zwar in affininvarianter Weise. Man kann daher z.B. eine zweite Fläche $\mathfrak{F}'$ betrachten, die dann in diesem Riemannschen Raume eingebettet erscheint. Dadurch ergeben sich neue Gesichtspunkte für die Geometrie der Flächenpaare und für die relative Differentialgeometrie, worauf in einer späteren Arbeit eingegangen werden soll.

---

[5]) CARTAN, E.: Les espaces de FINSLER. Paris 1934.

[6]) § 7 der vorliegenden Arbeit.

[7]) Die Krümmung der Eichfläche des Minkowskischen Raumes und die geometrische Deutung des einen Krümmungstensors des Finslerschen Raumes. Abh. math. Sem. Univ. Hamburg **20**, 41—51 (1955).

§ 1. Einführung der zu einer Hyperfläche gehörigen Raummetrik

Es sei $\mathfrak{F}$: $x^i = x^i(u^\alpha) \neq 0$ eine Hyperfläche in einem reellem Vektorraum, $\underset{0}{x}{}^i$ ein Punkt dieser Fläche. Dann wird es (bei gewissen Voraussetzungen über Differenzierbarkeit und lineare Unabhängigkeit) unter den Quadriken mit Mittelpunkt 0 genau eine, $\mathfrak{F}_2(\underset{0}{x})$, geben, welche $\mathfrak{F}$ in $\underset{0}{x}$ oskuliert. Diese Quadrik dient zu einer Einführung einer Riemannschen Maßbestimmung im Vektorraum (mit Ausnahme des Nullpunktes):

Ein im Punkte $r \cdot \underset{0}{x}{}^i$ $(r>0)$ angetragener Vektor y^i hat die Länge $L = L(y; \underset{0}{x})$, wenn folgendes gilt: Verschiebt man y^i parallel, so daß der Anfangspunkt in den Nullpunkt fällt, dann liegt sein Endpunkt auf der Fläche $L \cdot \mathfrak{F}_2(\underset{0}{x})$.

Damit ist für Kurven $x^i(t)$ (insbesondere für solche auf der Fläche $\mathfrak{F}$ selbst) ein Bogenelement definiert:

$$ds = L(\dot{x}^i; x^i)\, dt.$$

Wir werden die analytischen Voraussetzungen so einzurichten haben, daß

1. die oskulierende Quadrik eindeutig bestimmt ist,
2. die Flächenmetrik nicht ausgeartet ist.

Zur analytischen Durchführung der angegebenen Konstruktion denken wir uns das Flächenstück implizit gegeben durch

$$F(x^i) = 1, \quad F(r\,x^i) = r \cdot F(x^i) \qquad \text{für} \qquad r>0.$$

(Diese Darstellung ist immer möglich, wenn die Fläche sternförmig zum Nullpunkt liegt, d. h. mit jedem Halbstrahl höchstens einen Punkt gemeinsam hat.) Wir setzen zusätzlich voraus, daß die Funktion $F^2(x)$ viermal stetig differenzierbar sei (für $x \neq 0$). Es soll nun die Gleichung der oskulierenden Quadrik $\mathfrak{F}_2(\underset{0}{x})$ bestimmt werden. Wir nehmen an, die gesuchte Quadrik sei (in Abhängigkeit vom Flächenpunkt $\underset{0}{x}$) gegeben durch

$$g_{ik}(\underset{0}{x})\, y^i y^k = 1 \qquad (g_{ik} \text{ symmetrisch}).$$

$g_{ik}(\underset{0}{x})$ soll nun so bestimmt werden, daß

$$G(y) = F^2(y) - g_{ik}(\underset{0}{x})\, y^i y^k$$

für $y = \underset{0}{x}$ von möglichst hoher Ordnung verschwindet. Die Taylorsche Formel liefert (nach einfachen Umrechnungen mit Hilfe der Eulerschen Homogenitätsbeziehung)

$$G(y) = \left(\frac{1}{2}\frac{\partial^2 F^2(\underset{0}{x})}{\partial x^i \partial x^k} - g_{ik}(\underset{0}{x})\right) y^i y^k + R,$$

wobei R von mindestens dritter Ordnung in $y - \underset{0}{x}$ ist. Soll also $G(y)$ in $y = \underset{0}{x}$ von zweiter Ordnung verschwinden, so ergibt sich als eindeutige Festlegung

von $g_{ik}\left(\underset{0}{x}\right)$:

$$g_{ik}\left(\underset{0}{x}\right) = \frac{1}{2}\,\frac{\partial^2 F^2\left(\underset{0}{x}\right)}{\partial x^i \partial x^k}.$$

Damit ist die oskulierende Fläche zweiter Ordnung eindeutig bestimmt. Wiederholt man die Konstruktion für alle Flächenpunkte, und beachtet man, daß die Riemannsche Raummetrik nach der vorhin angegebenen geometrischen Konstruktion durch homogene Fortsetzung entsteht, so hat man allgemein für alle Raumpunkte x:

$$g_{ik}(x) = \frac{1}{2}\,\frac{\partial^2 F^2(x)}{\partial x^i \partial x^k}.$$

Die in der Fläche $\mathfrak{F}$ induzierte Metrik

$$g_{\alpha\beta} = g_{ik}\, x^i_\alpha\, x^k_\beta, \qquad x^i_\alpha = \frac{\partial x^i}{\partial u^\alpha}$$

(griechische Indizes für Flächenparameter, lateinische für Raumvektoren) ist für uns von besonderem Interesse.

Es sei bemerkt, daß der Tensor $g_{ik}(x)$ in der Minkowskischen Geometrie eine Rolle spielt: Wenn die Fläche konvex und geschlossen ist, und wenn $g_{ik}(x)\,y^i y^k$ für jedes x eine positiv definite Form ist (wenn also die oskulierenden Quadriken Ellipsoide sind), dann wird dieser Tensor als Fundamentaltensor der Minkowskischen Geometrie verwendet. Dieser Bezug zur Minkowskischen Geometrie wird uns später interessieren.

§ 2. Übereinstimmung der Flächenmetrik mit der der radialaffinen Geometrie von Salkowski[8])

SALKOWSKI (l. c.[2]) hat eine Maßbestimmung für Flächen im radialaffinen Raum definiert durch

$$\gamma_{\alpha\beta} = \frac{(x_{\alpha\beta}, x_1, \ldots, x_{n-1})}{(x, x_1, \ldots, x_{n-1})}.$$

[Die Determinante $(x, x_1, \ldots, x_{n-1})$ soll von Null verschieden sein; d.h. die Tangentialebene nicht durch den Nullpunkt gehen.] Wir wollen zeigen, daß für die im vorigen Abschnitt definierte Flächenmetrik $g_{\alpha\beta}$ gilt: $g_{\alpha\beta} = -\gamma_{\alpha\beta}$. Dazu leiten wir ein einfaches Lemma her:

Lemma. *Für jeden Flächenpunkt $x^i = x^i(u^\alpha)$ sind die beiden linearen Funktionale*

$$\lambda(y) = \frac{(y, x_1 \ldots x_{n-1})}{(x, x_1 \ldots x_{n-1})} \qquad \textit{und} \qquad \mu(y) = g_{ik}(x)\, x^i y^k$$

gleich:

$$g_{ik}(x)\, x^i y^k = \frac{(y, x_1, \ldots, x_{n-1})}{(x, x_1, \ldots, x_{n-1})}.$$

[8]) Man vergleiche dazu auch eine demnächst in den Publ. math. (Debrecen) erscheinende Note des Verf.: Eine Beziehung zwischen affiner und Minkowskischer Differentialgeometrie.

Beweis. Offensichtlich gilt

$$\lambda(x) = 1, \quad \lambda(x_\alpha) = 0.$$

Andererseits hat man auch $\mu(x) = g_{ik}(x)\, x^i x^k = 1$, und durch Differentiation nach u folgt daraus

$$0 = \frac{\partial}{\partial u^\alpha}\,(g_{ik}(x)\, x^i x^k) = 2 g_{ik}(x)\, x^i x_\alpha^k = 2 \cdot \mu(x_\alpha).$$

(Dabei wurde wieder von der Eulerschen Homogenitätsrelation Gebrauch gemacht:

$$\frac{\partial g_{ik}(x)}{\partial x^l}\, x^k = 0.)$$

Die beiden Linearformen λ und μ stimmen also für die n linear unabhängigen Vektoren x, x_α überein, woraus die Behauptung folgt.

Aus $g_{ik}(x)\, x^i x_\alpha^k = 0$ folgert man wieder unter Beachtung der Eulerschen Relation durch Differentiation nach u^β:

$$g_{ik}(x)\, x_\alpha^i x_\beta^k + g_{ik}(x)\, x_{\alpha\beta}^i x^k = 0,$$

und daraus sowie aus dem Lemma, angewendet für $y = x_{\alpha\beta}$, folgt

$$\gamma_{\alpha\beta} = g_{ik}(x)\, x^i x_{\alpha\beta}^k = -\, g_{ik}(x)\, x_\alpha^i x_\beta^k = -\, g_{\alpha\beta}.$$

Die affine Flächenmetrik $\gamma_{\alpha\beta}$ von SALKOWSKI *und unsere Metrik $g_{\alpha\beta}$ aus § 1 stimmen bis aufs Vorzeichen überein.*

Aus der geometrischen Einführung unserer Maßbestimmung in § 1 ergibt sich somit zugleich eine neue geometrische Deutung der Metrik von SALKOWSKI.

§ 3. Die Ableitungsgleichungen der affinen Flächentheorie und die kubische Fundamentalform

Da die Fläche $\mathfrak{F}$ gemäß § 1 in einem Riemannschen Raum eingebettet erscheint, kann man den Normalenvektor n^i und die zweite Fundamentalform $L_{\alpha\beta}$ der Fläche $\mathfrak{F}$ in diesem Riemannschen Raum bestimmen. Wir setzen jetzt voraus, daß $g_{ik}(x)$ invertierbar ist:

$$g^{ir} g_{rk} = \delta_k^i.$$

Wegen

$$g_{ik}(x)\, x^i x_\alpha^k = 0, \quad g_{ik}(x)\, x^i x^k = 1$$

kommt für den Normalvektor nur in Frage $n^i = \pm x^i$; wir entscheiden uns hier aus gleich ersichtlichen Gründen für das Minuszeichen:

$$n^i = -\, x^i.$$

Die zweite Fundamentalform berechnet sich zu (; für kovariante Ableitung)

$$L_{\alpha\beta} = x^i_\alpha x^k_\beta n_{k;i} = x^i_\alpha x^k_\beta (g_{kj}(x)\, x^j)_{;i}$$
$$= x^i_\alpha x^k_\beta g_{kj}\left(\delta^j_i + \left\{ \begin{matrix} j \\ i\,r \end{matrix} \right\} x^r\right) = g_{\alpha\beta}.$$

Es ergibt sich also:

Die zweite Fundamentalform fällt mit der ersten zusammen. (Alle Flächenpunkte sind Nabelpunkte; die Indikatrix teilt also auch diese Eigenschaft mit der Sphäre des euklidischen Raumes.)

Daher reduzieren sich die Ableitungsgleichungen auf die „Gaußsche" Gleichung:

$$x^i_{\alpha\beta} = A^\varrho_{\alpha\beta} x^i_\varrho + L_{\alpha\beta} n^i = A^\varrho_{\alpha\beta} x^i_\varrho - g_{\alpha\beta} x^i.$$

Nach Überschieben mit $g_{ik}(x)\, x^k_\gamma$ folgt daraus

$$g_{ik}(x)\, x^i_{\alpha\beta} x^k_\gamma = g_{\varrho\gamma} A^\varrho_{\alpha\beta}.$$

Um $A^\varrho_{\alpha\beta}$ zu berechnen, differenzieren wir die Beziehungen

$$g_{\alpha\beta} = g_{ik} x^i_\alpha x^k_\beta = - g_{ik} x^i x^k_{\alpha\beta}$$

nach u^γ:

$$\frac{\partial g_{\alpha\beta}}{\partial u^\gamma} = g_{ikj} x^i_\alpha x^k_\beta x^j_\gamma + g_{ik} x^i_{\alpha\gamma} x^k_\beta + g_{ik} x^i_\alpha x^k_{\beta\gamma} = - g_{ik} x^i_\gamma x^k_{\alpha\beta} - g_{ik} x^i x^k_{\alpha\beta\gamma}$$

(mit $g_{ikj} = \frac{\partial g_{ik}}{\partial x^j}$, Indizes an x^i für Ableitung). Aus der ersten Beziehung ergibt sich für die Christoffel-Symbole der Fläche:

$$\{\alpha\beta,\gamma\} = \frac{1}{2}\left(\frac{\partial g_{\alpha\gamma}}{\partial u^\beta} - \frac{\partial g_{\alpha\beta}}{\partial u^\gamma} + \frac{\partial g_{\beta\gamma}}{\partial u^\alpha}\right) = \frac{1}{2} g_{ikj} x^i_\alpha x^k_\beta x^j_\gamma + g_{ik} x^i_{\alpha\beta} x^k_\gamma.$$

Wir führen den folgenden Tensor ein

$$a_{\alpha\beta\gamma} = \tfrac{1}{2} g_{ikj} x^i_\alpha x^k_\beta x^j_\gamma; \qquad a^\varrho_{\alpha\beta} = g^{\varrho\gamma} a_{\alpha\beta\gamma}, \qquad g^{\varrho\gamma} g_{\gamma\sigma} = \delta^\varrho_\sigma$$

und erhalten

$$g_{\varrho\gamma} A^\varrho_{\alpha\beta} = g_{ik} x^i_{\alpha\beta} x^k_\gamma = \{\alpha\beta,\gamma\} - a_{\alpha\beta\gamma}.$$

Die Ableitungsgleichungen gewinnen damit die Gestalt:

$$x^i_{\alpha\beta} = \left(\left\{ \begin{matrix} \varrho \\ \beta\,\alpha \end{matrix} \right\} - a^\varrho_{\alpha\beta}\right) x^i_\varrho - g_{\alpha\beta} x^i.$$

Ihre Koeffizienten drücken sich allein durch $g_{\alpha\beta}$ und $a_{\alpha\beta\gamma}$ aus; es ist darum naheliegend, mittels des symmetrischen Tensors $a_{\alpha\beta\gamma}$ eine *kubische Fundamentalform* zusätzlich zur metrischen Fundamentalform $g_{\alpha\beta}$ einzuführen.

Schon aus der Gestalt der Ableitungsgleichungen geht hervor, daß die kubische Fundamentalform verantwortlich ist für die Abweichung des Raumes vom euklidischen Verhalten. Wir bemerken dazu noch:

$a_{\alpha\beta\gamma} = 0$ ist notwendig und hinreichend dafür, daß $g_{ijk} = 0$, d. h. daß die Fläche $\mathfrak{F}$ eine Mittelpunktsquadrik ist.

Dies ist in der Richtung „dann" selbstverständlich. In der Richtung „nur dann" folgt die Bemerkung aus den beiden Beziehungen

$$0 = a_{\alpha\beta\gamma} = g_{ijk}\, x^i_\alpha\, x^j_\beta\, x^k_\gamma \quad \text{und} \quad 0 = g_{ijk}\, x^i.$$

Wir zeigen nun noch, daß unsere kubische Fundamentalform a identisch ist mit der von SALKOWSKI eingeführten, deren Koeffizienten definiert sind durch

$$\bar{a}_{\alpha\beta\gamma} = \frac{(x^i_{\alpha;\beta;\gamma}, x_1, \ldots, x_{n-1})}{(x, x_1, \ldots, x_{n-})}.$$

Nach unserem Lemma (§ 2) ist zunächst

$$\bar{a}_{\alpha\beta\gamma} = g_{ik}\, x^i_{\alpha;\beta;\gamma}\, x^k.$$

Nun gilt nach den Ableitungsgleichungen

$$x^i_{\alpha;\beta} = x^i_{\alpha\beta} - \left\{ \begin{matrix} \varrho \\ \alpha\,\beta \end{matrix} \right\} x^i_\varrho = a^\varrho_{\alpha\beta}\, x^i_\varrho - g_{\alpha\beta}\, x^i.$$

Andererseits ist

$$0 = g_{\alpha\beta;\gamma} = -\left(g_{ik}\, x^i_{\alpha;\beta}\, x^k\right)_{;\gamma} = -g_{ik}\, x^i_{\alpha;\beta;\gamma}\, x^k - g_{ik}\, x^i_{\alpha;\beta}\, x^k_\gamma.$$

Aus beidem zusammen ergibt sich

$$\bar{a}_{\alpha\beta\gamma} = g_{ik}\, x^i_{\alpha;\beta;\gamma}\, x^k = -g_{ik}\, x^i_{\alpha;\beta}\, x^k_\gamma = g_{ik}\, x^i_\varrho\, x^k_\gamma\, a^\varrho_{\alpha\beta} = a_{\alpha\beta\gamma},$$

was zu beweisen war.

Es sei bemerkt, daß $a^\alpha_{\beta\gamma}$ außerdem übereinstimmt mit dem in VARGAS Arbeit (l. c. [7]) auftretenden Tensor $\sum^\alpha_{\beta\gamma}$.

§ 4. Die Integrabilitätsbedingungen und die Bestimmung einer Fläche durch die beiden Fundamentalformen

Um die Integrabilitätsbedingungen der Ableitungsgleichungen

$$x^i_{\alpha\beta} = \left(\left\{ \begin{matrix} \varrho \\ \alpha\,\beta \end{matrix} \right\} - a^\varrho_{\alpha\beta} \right) x^i_\varrho - g_{\alpha\beta}\, x^i$$

zu finden, berechnen wir $x^i_{\alpha\beta\gamma} - x^i_{\alpha\gamma\beta}$ und drücken alles vermöge der Ableitungsgleichungen durch x und x_ϱ aus. Man rechnet leicht nach, daß der Koeffizient von x hierin die Symmetrie der kubischen Fundamentalform ausdrückt. Das Verschwinden der Koeffizienten von x^i_ϱ ergibt jedoch nichttriviale Bedingungen:

$$R^\varrho_{\beta\gamma\delta} = g_{\beta\gamma}\,\delta^\varrho_\delta - g_{\beta\delta}\,\delta^\varrho_\gamma + a^\sigma_{\beta\gamma}\, a^\varrho_{\sigma\delta} - a^\sigma_{\beta\delta}\, a^\varrho_{\sigma\gamma} + g^{\varrho\sigma}\, a_{\sigma\beta\gamma;\delta} - g^{\varrho\sigma}\, a_{\sigma\beta\delta;\gamma}.$$

Dies sind die notwendigen und hinreichenden Integrabilitätsbedingungen, denen wir jedoch im Hinblick auf geometrische Anwendungen noch eine etwas andere Gestalt geben wollen.

Zunächst bestätigt man sofort durch direkte Rechnung, daß die kovariante Ableitung von $a_{\alpha\beta\gamma}$ symmetrisch in allen Indizes ist:

$$a_{\alpha\beta\gamma;\delta} = a_{\alpha\beta\delta;\gamma}. \tag{4.2}$$

(Die Rechnung ist im wesentlichen dieselbe wie bei O. VARGA, l. c.[7]), p. 46 — dort wird der $a_{\alpha\beta\gamma}$ entsprechende Tensor mit $\sum_{\alpha\beta\gamma}$ bezeichnet — und darf daher hier übergangen werden.)

Setzt man dies in die Integrabilitätsbedingungen (4.1) ein, so ergibt sich (wenn man $R_{\alpha\beta\gamma\delta}$ für den Krümmungstensor der Fläche schreibt):

$$R_{\alpha\beta\gamma\delta} = (g_{\alpha\delta}g_{\beta\gamma} - g_{\alpha\gamma}g_{\beta\delta}) + a^{\varrho}_{\beta\gamma}a_{\varrho\alpha\delta} - a^{\varrho}_{\beta\delta}a_{\varrho\alpha\gamma}. \tag{4.3}$$

Zusammen mit der Symmetrie der beiden Fundamentalformen $g_{\alpha\beta}$ und $a_{\alpha\beta\gamma}$ sind (4.2) und (4.3) die notwendigen und hinreichenden Integrabilitätsbedingungen der radialaffinen Flächentheorie. Es gilt dann der

Hauptsatz der radialaffinen Flächentheorie. *Zu vorgegebenen Funktionen $\tilde{g}_{\alpha\beta}(u^\delta)$, $\tilde{a}_{\alpha\beta\gamma}(u^\delta)$, $\tilde{g}$ nicht ausgeartet, $\tilde{g}$ und $\tilde{a}$ symmetrisch, welche den Bedingungen* (4.2) *und* (4.3) [*oder der etwas schwächeren Bedingung* (4.1)] *genügen, gibt es eine und bis auf homogene Affinitäten genau eine Fläche, deren Fundamentalformen $g_{\alpha\beta}$ und $a_{\alpha\beta\gamma}$ mit $\tilde{g}$ und $\tilde{a}$ übereinstimmen.*

Wenn die Integrabilitätsbedingungen erfüllt sind, gibt es zu vorgegebenen Anfangsbedingungen $x^i(\underset{0}{u}{}^\delta) = \underset{0}{x}{}^i$, $x^i_\alpha(\underset{0}{u}{}^\delta) = \underset{0}{x}{}^i_\alpha$ jedenfalls genau eine Lösung von

$$x^i_{\alpha\beta} = \left(\left\{\begin{matrix}\varrho\\ \alpha\beta\end{matrix}\right\} - \tilde{a}^{\varrho}_{\alpha\beta}\right)x^i_\varrho - \tilde{g}_{\alpha\beta}x^i.$$

Sind $\underset{0}{x}{}^i, \underset{0}{x}{}^i_\alpha$ linear unabhängig, so gilt das gleiche für die Lösung in einer Umgebung von $\underset{0}{u}{}^\delta$. Für die zu der dann durch $x^i(u^\delta)$ definierten Fläche gehörigen Fundamentalformen g, a gelten die Ableitungsgleichungen

$$x^i_{\alpha\beta} = \left(\left\{\begin{matrix}\varrho\\ \alpha\beta\end{matrix}\right\} - a^{\varrho}_{\alpha\beta}\right)x^i_\varrho - g_{\alpha\beta}x^i,$$

und Koeffizientenvergleich der beiden Systeme ergibt die Gleichheit von $\tilde{g}, \tilde{a}$ mit den Fundamentalformen der Fläche.

Sind durch $\underset{1}{x}{}^i, \underset{1}{x}{}^i_\alpha$ andere Anfangsbedingungen vorgeschrieben, so erhält man eine Lösung mit diesen Anfangswerten, wenn man auf die eben gefundene Lösung diejenige Affinität anwendet, die $\underset{0}{x}{}^i\ \underset{0}{x}{}^i_\alpha$ in $\underset{1}{x}{}^i, \underset{1}{x}{}^i_\alpha$ überführt; aus der Eindeutigkeit folgt der Hauptsatz.

§ 5. Krümmungen von Raum und Fläche

Da die Fläche $\mathfrak{F}$ als Untermannigfaltigkeit des Raumes mit der Riemannschen Metrik g_{ik} betrachtet werden kann, besteht zwischen dem Krümmungstensor R_{ijkl} des Raumes und dem Krümmungstensor $R_{\alpha\beta\gamma\delta}$ der Fläche be-

kanntlich (vgl. etwa SCHOUTEN, Ricci-Calculus, 2nd. ed., p. 242) der Zusammenhang

$$x^i_\alpha x^j_\beta x^k_\gamma x^l_\delta R_{ijkl} = R_{\alpha\beta\gamma\delta} + (g_{\alpha\gamma} g_{\beta\delta} - g_{\alpha\delta} g_{\beta\gamma}). \tag{5.1}$$

Aus der Integrabilitätsbedingung (4.3) ergibt sich

$$x^i_\alpha x^j_\beta x^k_\gamma x^l_\delta R_{ijkl} = a^\varrho_{\beta\gamma} a_{\varrho\alpha\delta} - a^\varrho_{\beta\delta} a_{\varrho\alpha\delta}. \tag{5.2}$$

An anderer Stelle[1]) ist der Krümmungstensor R_{ijkl} berechnet worden:

$$R_{ijkl} = -\tfrac{1}{4} g^{rs} (g_{irk} g_{jsl} - g_{irl} g_{jsk}). \tag{5.3}$$

Daraus folgt

$$R_{ijkl} x^i = R_{ijkl} x^k = 0 \tag{5.4}$$

nach der Eulerschen Homogenitätsrelation. Aus (5.2) und (5.4) zusammen ergibt sich:

Notwendig und hinreichend für $R_{ijkl} = 0$ *ist*

$$a^\varrho_{\beta\gamma} a_{\varrho\alpha\delta} - a^\varrho_{\beta\delta} a_{\varrho\alpha\gamma} = 0.$$

Aus (5.1) folgt außerdem — wie auch bereits VARGA bemerkt hat —:

Notwendig und hinreichend dafür, daß $R_{ijkl} \equiv 0$, ist, daß die Flächenkrümmung konstant gleich 1 ist:

Die Flächen konstanter Affinkrümmung 1 *sind dadurch gekennzeichnet, daß die Krümmung der zugehörigen Raummetrik identisch verschwindet.*

Schließlich ist noch zu bemerken, daß die Quadriken mit Mittelpunkt 0 konstante Krümmung 1 haben, da für sie die Raumkrümmung nach (5.3) wegen $g_{ijk} = 0$ verschwindet.

Da die zweite Fundamentalform gleich der ersten ist, ist die *Normalkrümmung einer Kurve auf der Fläche* $\mathfrak{F}$ *identisch gleich* 1.

Die bisherigen Betrachtungen dieses Abschnittes gelten naturgemäß nur für Räume unendlicher oder endlicher Dimension ≥ 3. Für Kurven in der radialaffinen Ebene hat sich bereits früher[1]) ergeben, daß *die Krümmung der zugehörigen Metrik in der Ebene identisch verschwindet.* Man rechnet leicht nach, daß für die Krümmung K der definierenden Kurve $\mathfrak{F}$ gilt: $K \equiv 1$.

§ 6. Spezielle Probleme

In diesem Abschnitt sollen einige speziellere Fragen untersucht werden, wobei wir uns — wie in der vorliegenden Arbeit überhaupt — auf die Betrachtung solcher geometrischer Größen beschränken wollen, die für die geometrische Interpretation von E. CARTANs Theorie der Finsler-Räume (§ 7) von Interesse sind.

Wir haben bereits gesehen, daß der zur Hyperfläche $\mathfrak{F}$ in der von ihr erzeugten Riemannschen Raummetrik senkrechte Vektor im wesentlichen

gleich dem Ortsvektor x^i ist, so daß auf diesem Weg keine neue geometrische Größe erhalten wird. Man kann aber (im Falle endlicher Raumdimension) durch

$$N^i = \frac{1}{n-1} g^{\alpha\beta} x^i_{\alpha;\beta} \tag{6.1}$$

einen „Affinnormalvektor" erklären, der bekanntlich interessante geometrische Eigenschaften hat (Tangente der Schwerlinie). Dieser Vektor soll durch die Raummetrik ausgedrückt werden. Nach den Ableitungsgleichungen ist

$$\begin{aligned} N^i &= \frac{1}{n-1} g^{\alpha\beta} a^\varrho_{\alpha\beta} x^i_\varrho - x^i = -x^i + \frac{1}{2(n-1)} g^{\alpha\beta} g^{\varrho\sigma} g_{jkl} x^j_\alpha x^k_\beta x^l_\varrho x^i_\sigma \\ &= \frac{1}{2(n-1)} (g^{lj} - x^l x^j)(g^{ik} - x^i x^k) g_{jkl} - x^i \end{aligned}$$

$$N^i = \frac{1}{2(n-1)} g^{lj} g^{ik} g_{jkl} - x^i. \tag{6.2}$$

Die *Affinsphären* sind diejenigen Flächen, deren sämtliche Affinnormalen durch 0 gehen; sie *sind also gekennzeichnet durch jede der beiden folgenden Relationen:*

$$g^{\varrho\sigma} a_{\varrho\sigma\beta} = 0, \tag{6.3}$$

$$g^{ik} g_{ikl} = \frac{\partial g}{\partial x^l} = 0 \qquad (g = \det g_{ik}). \tag{6.4}$$

Eine andere interessante Beziehung ergibt sich durch die Betrachtung der zu $\mathfrak{F}$ dualen Hyperfläche $\mathfrak{F}^*$ im Dualraum. (Diese Überlegungen gelten auch wieder im unendlichdimensionalen Fall.) Ordnet man vermöge

$$x_i = g_{ik}(x)\, x^k \tag{6.5}$$

jedem Punkte des Raumes einen Punkt des Dualraumes zu, so wird dabei gerade $\mathfrak{F}$ in $\mathfrak{F}^*$ übergeführt. Da $\mathfrak{F}^*$ gegeben ist durch

$$g^{ij}(x)\, x_i x_j = 1$$

und die g^{ij} bekanntlich die zweiten Ableitungen von $\frac{1}{2} H^2$ sind (H Minkowskische Stützfunktion von $\mathfrak{F}$), ist g^{ij} die zu $\mathfrak{F}^*$ gehörige Riemannsche Metrik des Dualraumes. Die Abbildung (6.5) ist wegen

$$d(x_i) = g_{ikl}(x)\, x^k dx^l + g_{ik}(x)\, x^k = g_{ik}(x^i)\, x^k$$

eine Isometrie dieser Riemannschen Metriken:

$$g^{ik} d(x_i)\, d(x_k) = g_{ik}\, dx^i dx^k;$$

speziell ergibt sich eine isometrische Abbildung zwischen $\mathfrak{F}$ und $\mathfrak{F}^*$, *deren radialaffine metrische Fundamentalformen also in entsprechenden Punkten übereinstimmen.*

Man kann den Dualraum vermöge einer nicht-ausgearteten homogenlinearen Abbildung A (a^{ij} symmetrisch) in den Ausgangsraum abbilden.

Dann sind $\mathfrak{F}$ und $A\,\mathfrak{F}^*$ in diesem Raum *isometrische Flächen*, welche übrigens polarreziprok bezüglich der Quadrik $a_{ik}\,y^i y^k = 1$ (a_{ik} Inverse zu a^{ij}) sind[9]). Man rechnet leicht nach, daß die kubischen Fundamentalformen von $\mathfrak{F}$ und $\mathfrak{F}^*$ sich nur im Vorzeichen unterscheiden.

§ 7. Affingeometrische Interpretation von E. CARTANs Geometrie der Finsler-Räume

Ein Finslerscher Raum ist eine differenzierbare Mannigfaltigkeit, in deren (radialaffinen) Tangentialräumen jeweils eine Hyperfläche $\mathfrak{F}$ als Indikatrix ausgezeichnet ist, deren Vektoren die Länge 1 zugeschrieben wird. Es ist daher von vornherein zu erwarten, daß die nur im Tangentialraum eines Punktes des Finslerschen Raumes definierten Größen eine affingeometrische Bedeutung besitzen[10]).

a) Die *Metrik* $g_{ik}(x)$ ist identisch mit der zur Indikatrix gehörigen *Riemannschen Metrik im Tangentialraum.*

b) Die Koeffizienten $C_j{}^i{}_k$ des *„euklidischen Zusammenhangs"* des Finslerschen Raumes sind die Koeffizienten des *Levi-Civitaschen Parallelismus der Riemannschen Metrik.*

c) Der Vektor $A^i(x) = \frac{1}{2}\,F(x)\,g^{ij}\,\frac{\partial g}{\partial x^j}$ hängt nach (6.2) in einfacher Weise mit dem *Affinnormalvektor* N^i zusammen:

$$A^i(x) = (n-1)\,F(x)\left(\frac{x^i}{F(x)} + N^i\left(\frac{x}{F(x)}\right)\right).$$

Insbesondere folgt daraus, daß $A^i = 0$ kennzeichnend ist für diejenigen Finsler-Räume, deren Indikatrizen Affinsphären sind, was bereits von A. DEICKE[11]) bemerkt und verwendet wurde. (Für die Definition von A^i: CARTAN, l. c. p. 13.)

d) Von den drei *Krümmungstensoren* des Finslerschen Raumes wird hier der erste, mit S_{ijkh} bezeichnete erfaßt (CARTAN p. 34, Formel XVI). Für ihn gilt:

$$S_{ijkh}(x) = \frac{1}{F^2(x)}\,R_{ijkh} \tag{7.1}$$

mit dem Krümmungstensor R_{ijkh} der Riemannschen Metrik[12]). Er verschwindet also für jeden zweidimensionalen Finsler-Raum identisch und ist

[9]) Daß bezüglich der „Einheitskugel" polarreziproke Flächen isometrisch sind, ist bereits in der Monographie von SALKOWSKI bemerkt (p. 128). Für konvexe Flächen $\mathfrak{F}$ hat E. R. LORCH die Isometrie der zu $\mathfrak{F}$ und $\mathfrak{F}^*$ gehörigen Raummetriken bewiesen: On the volume of smooth convex bodies in Hilbert space, Math. Z. **61**, 391—407 (1955).

[10]) Da wir nur einen einzigen festen Punkt des Finslerschen Raumes und dessen Tangentialraum betrachten, dürfen wir — um mit den bisherigen Bezeichnungen arbeiten zu können — von CARTANs Bezeichnung insofern abweichen, als wir für Tangentialvektoren x statt x' schreiben.

[11]) Die Finsler-Räume mit $A_i = 0$. Arch. d. Math. **4**, 45—51 (1953).

[12]) l. c. [1]) und Bull. Amer. Math. Soc. Abstract 62—3—357.

für höherdimensionale Räume dann und nur dann gleich Null, wenn die Indikatrizen konstante Affinkrümmung 1 haben.

Den aus (7.1) und (5.1) folgenden Zusammenhang mit der Riemannschen Krümmung der Indikatrix hat auch VARGA bemerkt (l. c. [7]).

e) Der Landsbergsche *Winkel* (*métrique angulaire* bei CARTAN, p. 13–14) $d\varphi$ zwischen zwei Richtungen, repräsentiert durch die Einheitsvektoren x^i und $x^i + dx^i$) ist *identisch mit der affinen Flächenmetrik* der Indikatrix:

$$d\varphi^2 = g_{ik}(x)\, dx^i dx^k.$$

Aus den Ergebnissen von § 5 und (7.1) folgt der bei CARTAN (p. 35) ausgesprochene und auch von VARGA bewiesene Satz, daß die Winkelmetrik dann und nur dann von konstanter Krümmung 1 ist, wenn $S_{ijkh} = 0$.

Die in § 6 gezeigte Isometrie zwischen $\mathfrak{F}$ und der Fläche $\mathfrak{F}^*$ im Dualraum (der „Figuratrix") läßt sich hier so interpretieren: Der Winkel zwischen zwei Hyperebenen, gemessen durch die Metrik des Dualraumes, ist gleich dem Winkel zwischen ihren Normalvektoren.

f) Die Finsler-Räume mit

$$A_{ijk|0} = \frac{\partial A_{ijk}}{\partial x^h}\, x^h = 0$$

haben nach CARTAN p. 37 die Eigenschaft, daß die Winkelmetrik bei Parallelverschiebung unverändert bleibt. Das heißt, daß die *Indikatrizen durch Affinverbiegung auseinander hervorgehen.*

Technische Hochschule München, Institut für Geometrie

(Eingegangen am 12. Oktober 1956)

NAUMANN, H.
Math. Zeitschr. Bd. 67, S. 75—82 (1957)

Über Vektorsterne und Parallelprojektionen regulärer Polytope

Von
HERBERT NAUMANN

Das Ziel dieser Abhandlung ist im wesentlichen eine n-dimensionale Verallgemeinerung des Satzes von K. POHLKE (1853), der besagt, daß jedes konvexe Sechseck mit Mittelpunkt im gewöhnlichen Raum als Parallelprojektion eines Würfels erhalten werden kann. Es wird gezeigt, daß jedes „Hyper-Zonoeder" (d.i. ein konvexes Polytop, dessen 2-dimensionale Begrenzungselemente zentrale Symmetrie besitzen) des n-dimensionalen Raumes R_n, das p verschiedene Kantenrichtungen aufweist, im R_d, wo $d = \max(p, 2n-1)$ ist, als Parallelprojektion des p-dimensionalen Würfels γ_p darstellbar ist[1]). Der Beweis dieses Satzes stützt sich auf Betrachtungen über Vektorsterne im Sinne von H. HADWIGER [2]. Weiter werden die Sätze über Vektorsterne auch auf Parallelprojektionen regulärer Simplizes α_n und regulärer 2^n-Zelle β_n angewandt.

§ 1. Gegeben seien in einem n-dimensionalen euklidischen Raum R_n p in einem festen Punkt — etwa dem Koordinatenursprung O — angreifende Vektoren $\mathfrak{a}_1, \mathfrak{a}_2, \ldots, \mathfrak{a}_p$, die den Raum R_n aufspannen mögen. Hierbei sei

$$\mathfrak{a}_i = (a_{i1}, a_{i2}, \ldots, a_{in}) \qquad (i = 1, \ldots, p). \tag{1}$$

Das System dieser Vektoren, das wir mit $\mathfrak{S}_p$ bezeichnen, heiße ein *Vektorstern*[2]). Die Punkte mit den Koordinaten $(a_{i1}, a_{i2}, \ldots, a_{in})$ $(i = 1, \ldots, p)$ mögen die *Endpunkte* von $\mathfrak{S}_p$ heißen. Wenn die Vektoren alle den gleichen Betrag $|\mathfrak{a}_1| = |\mathfrak{a}_2| = \cdots |\mathfrak{a}_p| = A$ haben und paarweise orthogonal sind, so nennen wir den Vektorstern im Falle $A = 1$ einen *Einheitsstern* $\mathfrak{E}_p$, bei beliebigem $A > 0$ einen um den Faktor A *dilatierten Einheitsstern* $\mathfrak{E}_p^{(A)}$.

Das skalare Produkt zweier Vektoren $\mathfrak{a}$ und $\mathfrak{b}$ sei im folgenden mit $\mathfrak{a}\mathfrak{b}$ bezeichnet. — Wir setzen $\mathfrak{x} = (x_1, x_2, \ldots, x_n)$ und führen den Richtungsvektor

$$\mathfrak{c} = (\cos\alpha_1, \cos\alpha_2, \ldots, \cos\alpha_n) = \frac{\mathfrak{x}}{|\mathfrak{x}|} = \frac{\mathfrak{x}}{R} \tag{2}$$

[1]) Wir verwenden hier für das reguläre Simplex des R_n, das reguläre 2^n-Zell und den n-dimensionalen Würfel die eingebürgerten Bezeichnungen α_n bzw. β_n bzw. γ_n (vgl. z. B. H. S. M. COXETER [1]). In der Abhandlung [3] haben wir leider für β_n die Bezeichnung „Maßpolytop" benutzt, die von P. H. SCHOUTE (Mehrdimensionale Geometrie, Bd. 2, Leipzig 1905) in einer anderen Bedeutung, nämlich für den n-dimensionalen Würfel γ_n gebraucht wurde.

[2]) Nach HADWIGER [2]. Bei COXETER ([1], S. 27 ff.) ist ein etwas anderer Sternbegriff eingeführt.

ein. Nun definieren wir das *„Trägheitsmoment"*[3]) des Vektorsterns $\mathfrak{S}_p$ bezüglich der zur Richtung $\mathfrak{c}$ orthogonalen Hyperebene $\mathfrak{c}\mathfrak{x}=0$ als den Ausdruck

$$J=\sum_{i=1}^{p}(\mathfrak{a}_i\mathfrak{c})^2. \tag{3}$$

Trägt man vom Koordinatenursprung aus in jeder Richtung $\mathfrak{c}$ den Vektor der Länge $R=\sqrt{J^{-1}}$ ab, wo J das zur Richtung $\mathfrak{c}$ gehörige Trägheitsmoment ist, so liegen die Endpunkte dieser Vektoren auf der Hyperfläche zweiter Ordnung

$$\sum_{k,l=1}^{n} J_{kl}\,x_k\,x_l=1 \quad \text{mit} \quad J_{kl}=\sum_{i=1}^{p} a_{ik}\,a_{il} \qquad (k,l=1,\ldots,n), \tag{4}$$

denn es ist

$$J=\sum_{i=1}^{p}\Big(\sum_{k=1}^{n} a_{ik}\cos\alpha_k\Big)^2=\sum_{k,l=1}^{n} J_{kl}\cos\alpha_k\cos\alpha_l=\sum_{k,l=1}^{n} J_{kl}\frac{x_k\,x_l}{R^2}=\frac{1}{R^2}.$$

Da die Vektoren von $\mathfrak{S}_p$ nach Voraussetzung nicht in einer Hyperebene liegen, so ist J nach Gl. (3) für jede Richtung $\mathfrak{c}$ positiv. Also ist $R=\sqrt{J^{-1}}$ für jede Richtung eine endliche, reelle Größe, d.h. durch (4) wird ein „Hyperellipsoid" dargestellt. Seiner Definition zufolge ist es dem gegebenen Vektorstern $\mathfrak{S}_p$ räumlich fest zugeordnet; wir nennen es das *„Trägheits-Hyperellipsoid"* von $\mathfrak{S}_p$. Wenn ein Vektorstern die besondere Eigenschaft besitzt, daß sein Trägheits-Hyperellipsoid eine Hyperkugel ist, d.h. wenn die n^2 Beziehungen

$$J_{kl}=\delta_{kl}J_{11} \qquad (k,l=1,\ldots,n) \tag{5}$$

erfüllt sind, so nennen wir den Vektorstern *halb-eutaktisch*; ist außerdem der Angriffspunkt des Vektorsterns gleich dem Schwerpunkt seiner Endpunkte, so nennen wir den Vektorstern *eutaktisch*[4]). Eine Menge von p Punkten nennen wir *halb-eutaktisch bezüglich* O, wenn das System der p Vektoren mit O als Angriffspunkt und den gegebenen Punkten als Endpunkten ein halb-eutaktischer Vektorstern ist; eine Menge von p Punkten nennen wir (schlechthin) *eutaktisch*, wenn sie halb-eutaktisch bezüglich des Schwerpunkts ist, d.h. wenn die gegebenen Punkte die Endpunkte eines eutaktischen Vektorsterns sind.

Aus der Eigenschaft quadratischer Matrizen, daß ihre Zeilenvektoren genau dann paarweise orthogonal und vom Betrag A sind, wenn ihre Spaltenvektoren paarweise orthogonal und vom Betrag A sind, folgt unmittelbar:

[3]) Vgl. [3], § 6.

[4]) Zu dem von SCHLÄFLI eingeführten Begriff „eutaktisch" siehe auch COXETER [1], S. 251ff., wo der Begriff auf eine etwas andere Weise eingeführt ist, aber — wie ein Vergleich mit Satz 3.2 zeigen wird — dasselbe besagt wie hier. — Im gewöhnlichen dreidimensionalen Raum besitzt der Begriff des halb-eutaktischen Vektorsterns eine relativ einfache anschauliche Bedeutung. Man zeigt nämlich leicht, daß im R_3 ein Vektorstern $\mathfrak{S}_p$ mit dem Angriffspunkt O genau dann halb-eutaktisch ist, wenn das auf O bezogene physikalische Trägheitsellipsoid seiner Endpunkte (bei Massenbelegung 1) eine Kugel ist.

Satz 1.1. *Ein Vektorstern $\mathfrak{S}_n$ im R_n ist genau dann ein dilatierter Einheitsstern $\mathfrak{E}_n^{(A)}$, wenn er halb-eutaktisch ist.*

Man zeigt leicht, daß durch senkrechte Parallelprojektion aus einem (halb-) eutaktischen Vektorstern stets wieder ein (halb-)eutaktischer Vektorstern entsteht. Da sich jede Matrix $((a_{ik}))$ vom Typ (p, n), deren Spaltenvektoren paarweise orthogonal und vom Betrag A sind, durch Hinzufügen von $p-n$ weiteren Spalten zum A-fachen einer orthogonalen Matrix erweitern läßt, so kann man jeden halb-eutaktischen Vektorstern $\mathfrak{S}_p$ des R_n $(p \geq n+1)$, dessen Vektoren durch Gl. (1) gegeben seien, als senkrechte Parallelprojektion des halb-eutaktischen Vektorsterns $\mathfrak{S}_p'$ des R_p mit den Vektoren

$$\mathfrak{a}_i' = (a_{i1}, a_{i2}, \ldots, a_{ip}) \qquad (i = 1, \ldots, p) \tag{6}$$

darstellen. Somit gilt unter Benutzung von Satz 1.1:

Satz 1.2. *Ein Vektorstern $\mathfrak{S}_p$ des R_n $(p \geq n+1)$ ist genau dann im R_p als senkrechte Parallelprojektion eines $\mathfrak{E}_p^{(A)}$ darstellbar, wenn er halb-eutaktisch ist.*

Sei nun der allgemeine Fall betrachtet. Durch eine orthogonale Transformation läßt sich stets das Trägheits-Hyperellipsoid eines gegebenen Vektorsterns auf Hauptachsen transformieren. Wir können also ohne Beschränkung der Allgemeinheit von jetzt an voraussetzen, daß der Vektorstern $\mathfrak{S}_p$ bereits in einem Koordinatensystem gegeben ist, in dem sein Trägheits-Hyperellipsoid in Hauptachsenform

$$\sum_{k=1}^{n} J_{kk} x_k^2 = 1 \quad \text{mit} \quad J_{kk} = \sum_{i=1}^{p} a_{ik}^2 > 0 \tag{7}$$

dargestellt wird. Es gilt dann also

$$J_{kl} = \sum_{i=1}^{p} a_{ik} a_{il} = 0 \qquad \text{für } k \neq l. \tag{8}$$

Sei

$$d_{ik} = \frac{a_{ik}}{\sqrt{J_{kk}}} \qquad (i = 1, \ldots, p,\ k = 1, \ldots, n). \tag{9}$$

Dann bilden die n Spaltenvektoren der Matrix $((d_{ik}))$ ein normiertes Orthogonalsystem; durch Hinzufügung von geeignet gewählten $p-n$ Spalten läßt sich die Matrix also zu einer quadratischen p-reihigen orthogonalen Matrix $((d_{ik}))$ erweitern, deren Elemente d_{ik} für $k = 1, \ldots, n$ der Beziehung (9) genügen. Im R_p wird sodann durch das System der Vektoren

$$\mathfrak{b}_i = (A\, d_{i1}, A\, d_{i2}, \ldots, A\, d_{ip}) \qquad (A > 0,\ i = 1, \ldots, p) \tag{10}$$

ein um den Faktor A dilatierter Einheitsstern $\mathfrak{E}_p^{(A)}$ gegeben. Falls $2n-1 > p$, so läßt sich zur weiteren Betrachtung $\mathfrak{E}_p^{(A)}$ durch Anhängen von $2n-p-1$ weiteren Komponenten vom Wert Null an die Komponenten seiner Vektoren in einen R_{2n-1} einbetten; für $p \geq 2n-1$ spielt sich die weitere Betrachtung

im R_p ab. — Durch Projektion von $\mathfrak{E}_p^{(A)}$ in der Richtung[5])

$$x_1 = \text{const},\ x_2 = \text{const}, \ldots, x_n = \text{const}$$

auf den n-dimensionalen Raum

$$(11) \qquad x_{n+k} = \begin{cases} \alpha_k \cdot x_k & \text{für } k = 1, \ldots, n-1 \\ 0 & \text{für } k = n, \ldots, p-n \end{cases} \quad (\textit{dies nur falls } p > 2n-1)$$

ergibt sich der Vektorstern der $(2n-1)$-dimensional dargestellten Vektoren[6])

$$(12) \quad \mathfrak{c}_i = (A\,d_{i1}, A\,d_{i2}, \ldots, A\,d_{in}, A\,\alpha_1 d_{i1}, A\,\alpha_2 d_{i2}, \ldots, A\,\alpha_{n-1} d_{i,n-1}) \quad (i = 1, \ldots, p),$$

der sich nach Ausführung der orthogonalen Transformation

$$(13) \qquad \left\{ \begin{array}{l} x_k' = \dfrac{1}{\sqrt{1+\alpha_k^2}}\, x_k + \dfrac{\alpha_k}{\sqrt{1+\alpha_k^2}}\, x_{n+k} \\ x_n' = x_n \\ x_{n+k}' = -\dfrac{\alpha_k}{\sqrt{1+\alpha_k^2}}\, x_k + \dfrac{1}{\sqrt{1+\alpha_k^2}}\, x_{n+k} \end{array} \right\} \quad (k = 1, \ldots, n-1)$$

durch die Vektoren

$$(14) \quad \mathfrak{c}_i' = \left(A\sqrt{1+\alpha_1^2}\,d_{i1}, A\sqrt{1+\alpha_2^2}\,d_{i2}, \ldots, A\sqrt{1+\alpha_{n-1}^2}\,d_{i,n-1}, A\,d_{in}, 0, \ldots, 0\right)$$

darstellen läßt. Für

$$(15) \qquad \left\{ \begin{array}{l} J_{nn} = \min(J_{11}, J_{22}, \ldots, J_{nn})\ ^{7}), \quad A = \sqrt{J_{nn}}, \\ \alpha_k = \sqrt{\dfrac{J_{kk} - J_{nn}}{J_{nn}}} \qquad (k = 1, \ldots, n-1) \end{array} \right.$$

ist dies im $(x_1', x_2', \ldots, x_n')$-Raum wegen (9) der durch (1) gegebene Vektorstern $\mathfrak{S}_p$.

Damit ist gezeigt:

Satz 1.3. *Jeder Vektorstern $\mathfrak{S}_p$ des R_n ist im R_d, wo $d = \max(p, 2n-1)$, als Parallelprojektion eines dilatierten Einheitssterns $\mathfrak{E}_p^{(A)}$ des R_p darstellbar.*

§ 2. Da die Endpunkte eines $\mathfrak{E}_p^{(A)}$ des R_p die Ecken eines regulären Simplex α_{p-1} des R_{p-1} sind, so kann man mittels Satz 1.3 beweisen, daß sich jede Menge von p Punkten des R_n, die den R_n aufspannen, im R_d als Parallelprojektion der Ecken eines α_{p-1} darstellen läßt, wo $d = \max(p, 2n-1)$. Die folgende Betrachtung liefert ein Kriterium für die Möglichkeit einer Darstellung durch orthogonale Parallelprojektion eines α_{p-1} im R_{p-1} und führt sodann, im allgemeinen Fall, für die genannte Zahl d zu einem etwas günstigeren Wert.

---

[5]) Der im Uneigentlichen gelegene „projizierende Raum" ist also im Falle $2n-1 > p$ ein $(n-2)$-dimensionaler Raum, im Falle $p \geq 2n-1$ ein $(p-n-1)$-dimensionaler Raum.

[6]) Die für $p > 2n-1$ im Anschluß an die $(2n-1)$-te Komponente folgenden Komponenten sind hier, da sie sämtlich gleich Null werden, einfach weggelassen.

[7]) Dies darf o.B.d.A. vorausgesetzt werden.

Zunächst führen wir weitere Begriffe ein. Ein Vektorstern, dessen Angriffspunkt O mit dem Schwerpunkt seiner Endpunkte zusammenfällt, heiße ein *zentrierter Vektorstern*. Ein Vektorstern $\mathfrak{S}_{n+1}$ des R_n, dessen Endpunkte den R_n aufspannen (d.h. dessen Endpunkte die Ecken eines Simplex sind), heiße ein *Simplexstern*. Ein zentrierter Vektorstern $\mathfrak{S}_{n+1}$ des R_n ist immer ein Simplexstern. Ein zentrierter Vektorstern, dessen Endpunkte die Ecken eines regulären Simplex α_n bilden, heiße ein *regulärer Simplexstern*.

Satz 2.1. *Ein Simplexstern ist genau dann regulär, wenn er eutaktisch ist.*

Sei ein Simplexstern $\mathfrak{S}_p$ des R_{p-1}, der von den Vektoren

$$\mathfrak{a}_i = (a_{i1}, a_{i2}, \ldots, a_{i,p-1}) \qquad (i = 1, \ldots, p) \tag{1}$$

gebildet werde, eutaktisch. Dann gilt also

$$\left.\begin{aligned} &\sum_{i=1}^{p} a_{ik} = 0 \qquad (2)\\ &\sum_{i=1}^{p} a_{ik} a_{il} = \delta_{kl} \sum_{i=1}^{p} a_{i1}^2 = \delta_{kl} J_{11} \qquad (3)\end{aligned}\right\} \quad (k, l = 1, \ldots, p-1).$$

Setzt man

$$a_{ip} = \sqrt{\frac{J_{11}}{p}} \qquad (i = 1, \ldots, p), \tag{4}$$

so erweist sich [unter wesentlicher Benutzung der Schwerpunktsbedingung (2)] die erhaltene p-reihige quadratische Matrix $((a_{ik}))$ als das $\sqrt{J_{11}}$-fache einer orthogonalen Matrix. Also bilden im R_p die Vektoren

$$\left(a_{i1}, a_{i2}, \ldots, a_{i,p-1}, \sqrt{\frac{J_{11}}{p}}\right) \qquad (i = 1, \ldots, p) \tag{5}$$

einen $\mathfrak{E}_p^{(A)}$ mit $A = \sqrt{J_{11}}$. Damit ist aber klar, daß die Endpunkte von $\mathfrak{S}_p$ die Ecken eines regulären Simplex α_{p-1} der Kantenlänge $\sqrt{2 J_{11}}$ repräsentieren.

Sei nun umgekehrt durch Gl. (1) ein regulärer und folglich zentrierter Simplexstern $\mathfrak{S}_p$ gegeben, dessen Endpunkte die Ecken eines α_{p-1} der Kantenlänge $\sqrt{2} \cdot A$ sind. Dieses Simplex ist im R_p durch die Endpunkte eines $\mathfrak{E}_p^{(A)}$ mit dem Angriffspunkt O darstellbar, wobei bekanntlich das α_{p-1} von O den senkrechten Abstand $A/\sqrt{p}$ besitzt und das Lot von O auf das α_{p-1} dieses im Schwerpunkt trifft. Damit läßt sich schließen, daß im R_p durch die Vektoren

$$\mathfrak{a}_i = \left(a_{i1}, a_{i2}, \ldots, a_{i,p-1}, \frac{A}{\sqrt{p}}\right) \qquad (i = 1, \ldots, p) \tag{6}$$

ein $\mathfrak{E}_p^{(A)}$ gegeben ist. Folglich ist, als Folge von Satz 1.1, der gegebene reguläre Simplexstern eutaktisch und ferner $A = \sqrt{J_{11}}$.

Folgerung aus Satz 2.1. *Ein Simplex ist genau dann regulär, wenn die Menge seiner Eckpunkte eutaktisch ist.*

Satz 2.2. *Ein Vektorstern $\mathfrak{S}_p$ des R_n $(p \geq n+2)$ ist genau dann im R_{p-1} als senkrechte Parallelprojektion eines regulären Simplexsterns darstellbar, wenn er eutaktisch ist.*

Sei durch

$$\mathfrak{a}_i = (a_{i1}, a_{i2}, \ldots, a_{in}) \qquad (i = 1, \ldots, p) \tag{7}$$

ein eutaktischer Vektorstern $\mathfrak{S}_p$ gegeben. Dann läßt sich die Matrix $((a_{ik}))$ durch Hinzufügen von $p-n$ Spalten zu dem $\sqrt{J_{11}}$-fachen einer orthogonalen Matrix erweitern. Da die Ausgangsmatrix nach Voraussetzung die Eigenschaft hat, daß jede ihrer Spalten die Komponentensumme Null ergibt, so können die Elemente der letzten (p-ten) Spalte der Erweiterungsmatrix untereinander gleich gewählt werden. Dann ergeben zwangsläufig alle übrigen Spalten die Komponentensumme Null. Betrachtet man nun im R_{p-1} den von den Vektoren $\mathfrak{a}_i' = (a_{i1}, a_{i2}, \ldots, a_{i,p-1})$ gebildeten Vektorstern, so erkennt man, daß $\mathfrak{S}_p$ senkrechte Parallelprojektion eines eutaktischen Vektorsterns $\mathfrak{S}_p'$ im R_{p-1}, d.h. nach Satz 2.1 eines regulären Simplexsterns ist.

Der zweite Teil des Beweises folgt einfach aus der Bemerkung, daß jede senkrechte Parallelprojektion eines eutaktischen Vektorsterns wieder einen eutaktischen Vektorstern ergibt.

Folgerung aus Satz 2.2. *Eine Menge von p Punkten des R_n $(p \geq n+2)$ ist genau dann im R_{p-1} als senkrechte Parallelprojektion der Ecken eines α_{p-1} darstellbar, wenn sie eutaktisch ist.*

Satz 2.3. *Jeder zentrierte Vektorstern $\mathfrak{S}_p$ des R_n ist im R_d, wo $d = \max(p-1, 2n-1)$, als Parallelprojektion eines regulären Simplexsterns darstellbar.*

Der Beweis verläuft im wesentlichen genau so wie in § 1 der Beweis zu Satz 1.3, nur braucht die mittels § 1, Gl. (9) eingeführte Matrix $((d_{ik}))$ vom Typ (p, n) nur um $p-n-1$ Spalten erweitert werden, so daß die $p-1$ Spaltenvektoren der entstandenen Matrix paarweise orthogonale Einheitsvektoren mit der Eigenschaft sind, daß ihre Komponentensumme jeweils verschwindet. (Zur Realisierung der letzten Bedingung vgl. den Beweis zum vorhergehenden Satz 2.2.) Demgemäß tritt im weiteren Verlauf des Beweises die Formel $d = \max(p-1, 2n-1)$ an die Stelle der Formel $d = \max(p, 2n-1)$ von § 1.

Beim Übergang zur konvexen Hülle der Endpunkte von $\mathfrak{S}_p$ ergibt sich, wenn man der Einfachheit halber annimmt, daß alle p Punkte Eckpunkte ihrer konvexen Hülle sind:

Satz 2.4. *Jedes konvexe Polytop des R_n mit p Ecken ist im R_d, wo $d = \max(p-1, 2n-1)$, als Parallelprojektion eines α_{p-1} darstellbar.*

Für $n=2$, $p=4$ ergibt sich speziell, daß jedes konvexe Viereck im gewöhnlichen Raum als Parallelprojektion eines regulären Tetraeders darstellbar ist.

§ 3. Ein Vektorstern ist *zentralsymmetrisch*, wenn mit jedem Vektor $\mathfrak{a}_i$ seiner Vektoren auch sein entgegengesetzter Vektor $-\mathfrak{a}_i$ zu ihm gehört.

Ist $\mathfrak{S}_p$ ein beliebiger Vektorstern, so bezeichnen wir denjenigen Vektorstern, dessen Vektoren zu den Vektoren von $\mathfrak{S}_p$ entgegengesetzt sind, mit $\mathfrak{S}_p^*$. Der Vektorstern, der aus den Vektoren von $\mathfrak{S}_p$ und $\mathfrak{S}_p^*$ (unter Berücksichtigung eventueller Vielfachheiten) besteht, sei mit $\{\mathfrak{S}_p, \mathfrak{S}_p^*\}$ bezeichnet. Mit dieser Schreibweise ist also jeder zentralsymmetrische Vektorstern, der von $2p$ Vektoren gebildet wird, in der Form $\{\mathfrak{S}_p, \mathfrak{S}_p^*\}$ darstellbar. Aus der Definition eines eutaktischen Vektorsterns folgt, daß $\{\mathfrak{S}_p, \mathfrak{S}_p^*\}$ genau dann eutaktisch ist, wenn $\mathfrak{S}_p$ halb-eutaktisch ist. Es ist ferner leicht einzusehen, daß $\{\mathfrak{S}_p, \mathfrak{S}_p^*\}$ genau dann als (senkrechte) Parallelprojektion eines $\{\mathfrak{E}_p^{(A)}, \mathfrak{E}_p^{(A)*}\}$ darstellbar ist, wenn $\mathfrak{S}_p$ als (senkrechte) Parallelprojektion eines $\mathfrak{E}_p^{(A)}$ darstellbar ist. Da schließlich die Endpunkte eines $\{\mathfrak{E}_p^{(A)}, \mathfrak{E}_p^{(A)*}\}$ die Ecken eines β_p bilden, so gelten die Sätze:

Satz 3.1. *Ein zentralsymmetrischer Vektorstern $\mathfrak{S}_{2n}$ des R_n ist genau dann ein $\{\mathfrak{E}_p^{(A)}, \mathfrak{E}_p^{(A)*}\}$, wenn er eutaktisch ist.*

Satz 3.2. *Ein zentralsymmetrischer Vektorstern $\mathfrak{S}_{2p}$ des R_n $(p \geq n+1)$ ist genau dann im R_p als senkrechte Parallelprojektion eines $\{\mathfrak{E}_p^{(A)}, \mathfrak{E}_p^{(A)*}\}$ darstellbar, wenn $\mathfrak{S}_{2p}$ eutaktisch ist.*

Folgerung aus Satz 3.2. *Ein konvexes Mittelpunktspolytop des R_n mit $2p$ Ecken $(p \geq n+1)$ ist genau dann im R_p als senkrechte Parallelprojektion eines β_p darstellbar, wenn die Menge der $2p$ Ecken eutaktisch ist.*

Satz 3.3. *Jedes konvexe Mittelpunktspolytop des R_n mit $2p$ Ecken ist im R_d, wo $d = \max(p, 2n-1)$, als Parallelprojektion eines β_p darstellbar.*

Für $n=2$, $p=3$ bzw. $p=2$ besagt dieser letztere Satz speziell, daß jedes konvexe Sechseck mit Mittelpunkt, ebenso jedes Parallelogramm im gewöhnlichen Raum als Parallelprojektion eines regulären Oktaeders aufgefaßt werden kann.

Es sei noch bemerkt, daß jeder zentralsymmetrische Vektorstern $\mathfrak{S}_{2p}$ stets ein zentrierter Vektorstern ist. Dies erlaubt die Anwendung der Sätze von § 2 auf zentralsymmetrische Vektorsterne bzw. auf konvexe Polytope mit Mittelpunkt. Speziell liegt hierin die bekannte Tatsache eingeschlossen, daß jedes β_p als senkrechte Parallelprojektion eines α_{2p-1} darstellbar ist.

§ 4. Ein konvexes Polytop des R_n, dessen 2-dimensionale Begrenzungselemente zentralsymmetrisch sind, heiße ein *Hyper-Zonoeder*[8]. Für $n=2$ ergeben sich trivialerweise die konvexen Mittelpunktspolygone, für $n=3$ sind es die Zonoeder. Wie in [1], S. 27 und 28 gezeigt wird, entsteht das allgemeine Zonoeder aus einem Vektorstern, der von den Vektoren $\mathfrak{a}_1, \mathfrak{a}_2, \ldots, \mathfrak{a}_p$ gebildet sei, indem man die konvexe Hülle der Punkte

$$\mu_1 \mathfrak{a}_1 + \mu_2 \mathfrak{a}_2 + \cdots + \mu_p \mathfrak{a}_p \qquad (\mu_i = 0 \text{ oder } 1) \tag{1}$$

bildet. Nimmt man der Einfachheit halber an, daß die $\mathfrak{a}_i$ lauter verschiedene Richtung haben, so entsprechen die $\mathfrak{a}_i$ genau den p verschiedenen Kantenrichtungen des abgeleiteten Zonoeders. Die Überlegungen lassen sich leicht

[8]) Vgl. COXETER [1], S. 257. Für $n=3$ (Zonoeder) vgl. a. a. O. S. 28ff.

auf Hyper-Zonoeder beliebiger Dimension übertragen. Wenn der Vektorstern ein $\mathfrak{E}_n^{(A)}$ ist, so ist, wie man unmittelbar einsieht, das in der obigen Weise abgeleitete Hyper-Zonoeder ein n-dimensionaler Würfel γ_n der Kantenlänge A. Sodann gilt: Eine Parallelprojektion, die einen $\mathfrak{E}_p^{(A)}$ auf den Vektorstern $\mathfrak{S}_p$ abbildet, bildet den aus dem $\mathfrak{E}_p^{(A)}$ abgeleiteten Würfel γ_n auf das aus dem $\mathfrak{S}_p$ abgeleitete Hyper-Zonoeder ab.

Damit lassen sich aus den Sätzen des § 1 die folgenden Sätze ableiten:

Satz 4.1[9]. *Ein Hyper-Zonoeder mit p verschiedenen Kantenrichtungen ist genau dann als senkrechte Parallelprojektion eines γ_p darstellbar, wenn es sich aus einem halb-eutaktischen Vektorstern ableiten läßt.*

Satz 4.2. *Jedes Hyper-Zonoeder des R_n, das p verschiedene Kantenrichtungen aufweist, ist im R_d, wo $d = \max(p, 2n-1)$, als Parallelprojektion eines γ_p darstellbar.*

Für $n=2$, $p=3$ ist das die klassische Aussage des Satzes von POHLKE.

Literatur

[1] COXETER, H. S. M.: Regular Polytopes. London 1948. — [2] HADWIGER, H.: Über ausgezeichnete Vektorsterne und reguläre Polytope. Comm. Math. Helv. **13**, 90—107 (1940). — [3] NAUMANN, H.: Beliebige konvexe Polytope als Schnitte und Projektionen höherdimensionaler Würfel, Simplizes und Maßpolytope. Math. Z. **65**, 91—103 (1956).

[9] Vgl. COXETER [1], S. 255.

Schwelm (Westf.), Tannenstr. 8

(Eingegangen am 16. Oktober 1956)

JACOBS, K.
Math. Zeitschr. Bd. 67, S. 83—92 (1957)

Fastperiodizitätseigenschaften allgemeiner Halbgruppen in Banach-Räumen

Von

KONRAD JACOBS

Einleitung

In einer bekannten Abhandlung haben L. ALAOGLU und G. BIRKHOFF folgenden Satz bewiesen.

Allgemeiner Ergodensatz. *Es sei* $\mathfrak{H} = \{h, \ldots\}$ *ein Banach-Raum und* $\mathfrak{G} = \{x, \ldots\}$ *eine beliebige Halbgruppe nichtdehnender Transformationen in* $\mathfrak{H}$. *Sowohl der Raum* $\mathfrak{H}$ *als auch sein dualer Raum* $\mathfrak{H}'$ *seien gleichmäßig-konvex. Dann enthält für jedes* $h \in \mathfrak{H}$ *die konvexe abgeschlossene Hülle* $\mathfrak{K}(h)$ *der „Bahn"* $\mathfrak{G}h = \{xh | x \in \mathfrak{G}\}$ *genau einen* $\mathfrak{G}$*-Fixpunkt* h_0.

Es wurden noch einige Abschwächungen der Voraussetzung erreicht und — nach Formulierung eines geeigneten Konvergenzbegriffs — Aussagen über die Approximation des „Mittelwertes" h_0 durch „Näherungsmittelwerte" $\frac{1}{n}\sum_{\nu=1}^{n} x_\nu h$ gemacht (vgl. G. BIRKHOFF [2], ALAOGLU-BIRKHOFF [1], RIESZ-NAGY [22]).

Man kann die Aussage des Ergodensatzes auch unter anderen Voraussetzungen beweisen, nämlich, wenn $\mathfrak{H}$ ein beliebiger *reflexiver* Banach-Raum und $\mathfrak{G}$ eine *abelsche* Halbgruppe von linearen Transformationen in $\mathfrak{H}$, die eine gemeinsame Schranke besitzen, ist (RIESZ [19], [20], [21], LORCH [14], YOSIDA-KAKUTANI [23], EBERLEIN [9], BOURBAKI [3], JACOBS [12], RIESZ-NAGY [22]). Bei DUNFORD-MILLER [8], DAY [5], [7] und DIXMIER [24] finden sich weitergehende Verallgemeinerungen in dieser Richtung.

Unter den eben genannten Voraussetzungen läßt sich der *Mittelwertaussage* des Ergodensatzes folgende *Fastperiodizitätsaussage* zur Seite stellen: $\mathfrak{H}$ zerfällt direkt in zwei $\mathfrak{G}$-invariante abgeschlossene lineare Teilräume $\mathfrak{P}$ und $\mathfrak{F}$; $\mathfrak{P}$ ist der Raum der unter $\mathfrak{G}$ fastperiodischen Vektoren, $\mathfrak{F}$ ist der Raum der Fluchtvektoren bezüglich $\mathfrak{G}$ (JACOBS [12]). Aus dieser Aussage kann die Ergodenaussage gefolgert werden.

In der vorliegenden Abhandlung soll diese Aussage unter den (noch etwas abgeschwächten) Voraussetzungen des Allgemeinen Ergodensatzes von ALAOGLU-BIRKHOFF bewiesen werden. Neben den Ergodensatz tritt damit der

Aufspaltungssatz. *Es sei* $\mathfrak{H}$ *ein Banachraum und* $\mathfrak{G}$ *eine beliebige Halbgruppe von linearen Transformationen in* $\mathfrak{H}$. *Sowohl* $\mathfrak{H}$ *als auch der duale Raum* $\mathfrak{H}'$ *seien (je für sich) in solcher Weise stark-konvex normierbar, daß* $\mathfrak{G}$

bzw. die duale Halbgruppe $\mathfrak{G}'$ nicht-dehnend ist; mindestens einer der beiden Räume sei unter derselben Nebenbedingung gleichmäßig-konvex normierbar. Dann zerfällt $\mathfrak{H}$ direkt in den Raum $\mathfrak{P}$ der $\mathfrak{G}$-fastperiodischen Vektoren ([12], Definition II.1) und den Raum $\mathfrak{F}$ der $\mathfrak{G}$-Fluchtvektoren ([12], Definition II.2), genauer: Jeder Vektor $h \in \mathfrak{H}$ besitzt genau eine Darstellung

$$h = p + f, \quad p \in \mathfrak{P},\ f \in \mathfrak{F},$$

deren Komponenten p, f linear, stetig und $\mathfrak{G}$-invariant von h abhängen (Satz 3.1 [Aufspaltungssatz II]).

Der Beweisgang hat viele Einzelheiten mit dem in [*12*] angegebenen Beweisgang gemeinsam; wo irgend möglich, soll daher auf [*12*] verwiesen werden. Die wesentlichsten Abweichungen treten im Beweise des Satzes 1.1 und im Satz 2.1 auf.

Genau wie in [*12*] ergibt sich ein neuer Beweis des Allgemeinen Ergodensatzes (Satz 2.3) unter etwas abgeschwächten Voraussetzungen. Man kann den Allgemeinen Ergodensatz leicht aus dem Aufspaltungssatz folgern, indem man den Mittelwertsatz für fastperiodische Funktionen (MAAK [*16*], KEINER [*13*]) auf unsere fastperiodischen Vektoren überträgt und zugleich berücksichtigt, daß alle Fluchtvektoren nach dem Satz von MAZUR ([*12*], Satz I.1) nur den Fixpunkt $0 \in \mathfrak{H}$ liefern. Man kann den Aufspaltungssatz als eine Erläuterung zum Allgemeinen Ergodensatz benützen. In § 6 wird etwas näher untersucht, welche Grenzen dem Ergodizitätscharakter von $\mathfrak{G}$ in $\mathfrak{H}$ durch die Nichtkommutativität von $\mathfrak{G}$ im allgemeinen gesetzt sind.

§ 1. Reversible Vektoren und Fluchtvektoren. Formulierung des Aufspaltungssatzes I

Wir betrachten eine *beliebige Halbgruppe* $\mathfrak{G}$ von stetigen linearen Transformationen $x, \ldots$ in einem *reflexiven* Banach-Raum ([*12*], Kap. I) $\mathfrak{H} = \{h, \ldots\}$ und nehmen an, daß eine Zahl $\Gamma \geq 1$ mit

$$\|xh\| \leq \Gamma \|h\| \qquad (x \in \mathfrak{G},\ h \in \mathfrak{H}),$$

d.h. eine gemeinsame Schranke der $x \in \mathfrak{G}$ existiert ($\|h\|$ bezeichne die in $\mathfrak{H}$ erklärte Norm). $\mathfrak{G}$ enthalte die identische Transformation e.

Für jedes $h \in \mathfrak{H}$ bezeichne $\mathfrak{G}(h)$ die *schwach-abgeschlossene Hülle* der Menge $\mathfrak{G}h = \{xh \mid x \in \mathfrak{G}\}$ (vgl. [*12*], Definition II.2).

Definition. *Ein Vektor $f \in \mathfrak{H}$ heißt ein Fluchtvektor (unter $\mathfrak{G}$), wenn $O \in \mathfrak{G}(f)$ gilt; die Gesamtheit aller $\mathfrak{G}$-Fluchtvektoren bezeichnen wir mit $\mathfrak{F}$. Ein Vektor $r \in \mathfrak{H}$ heißt ein $\mathfrak{G}$-reversibler Vektor, wenn $\mathfrak{G}(h) = \mathfrak{G}(r)$ für jedes $h \in \mathfrak{G}(r)$ gilt; die Gesamtheit aller $\mathfrak{G}$-reversiblen Vektoren bezeichnen wir mit $\mathfrak{R}$ (vgl. [12], Definitionen II.2 und III.2).*

Wir definieren die *schwache Hülle* $\overline{\mathfrak{T}}$ einer Menge $\mathfrak{T}$ von linearen Transformationen in $\mathfrak{H}$ wie in [*12*]. Es gilt:

1. Ist $\mathfrak{T}$ eine Halbgruppe, so ist auch $\overline{\mathfrak{T}}$ eine Halbgruppe. Ist dabei $\mathfrak{T}$ abelsch, so ist auch $\overline{\mathfrak{T}}$ abelsch.

2. Ist Γ eine gemeinsame Schranke der $x \in \mathfrak{T}$, so ist Γ auch eine gemeinsame Schranke der $\bar{x} \in \overline{\mathfrak{T}}$.

3. $\mathfrak{G}(h) = \overline{\mathfrak{G}} h = \{\bar{x} h | \bar{x} \in \overline{\mathfrak{G}}\}$ $(h \in \mathfrak{H})$.

4. Jeder unter $\mathfrak{T}$ invariante abgeschlossene lineare Teilraum von $\mathfrak{H}$ ist auch unter $\overline{\mathfrak{T}}$ invariant.

5. Es gilt $f \in \mathfrak{F}$ dann und nur dann, wenn $\bar{x} \in \overline{\mathfrak{G}}$ mit $\bar{x} f = 0 \in \mathfrak{H}$ existiert. Es gilt $r \in \mathfrak{R}$ dann und nur dann, wenn es zu jedem $\bar{x} \in \overline{\mathfrak{G}}$ ein $\bar{y} \in \overline{\mathfrak{G}}$ mit $\bar{y} \bar{x} r = r$ gibt (vgl. JACOBS [12]).

Man beweist unschwer den

Satz 1.1. *Jeder Vektor $h \in \mathfrak{H}$ besitzt mindestens eine Zerlegung*

$$h = r + f, \quad r \in \mathfrak{R}, \ f \in \mathfrak{F}. \tag{1}$$

Beweis. Es sei $h \in \mathfrak{H}$ beliebig gewählt. Genau wie in [12] zeigt man mittels des *Zornschen Lemmas:* Es gibt ein $\bar{x} \in \overline{\mathfrak{G}}$ mit $\bar{x} h \in \mathfrak{R}$. Es sei $\mathfrak{X} = \{\bar{x}^\nu | \nu = 0, 1, \ldots\}$. Dann ist $\overline{\mathfrak{X}}$ eine *abelsche* Unterhalbgruppe von $\overline{\mathfrak{G}}$. Wiederholt man die zuerst bezüglich $\mathfrak{G}$ angestellten Überlegungen, indem man $\mathfrak{X}$ an die Stelle von $\mathfrak{G}$ treten läßt, so gewinnt man ein $\bar{a} \in \overline{\mathfrak{X}}$, derart, daß $\bar{a} \bar{x} h = g$ $\mathfrak{X}$-reversibel ist. Nach Nr. 5 gibt es ein $\bar{b} \in \overline{\mathfrak{X}}$ mit $\bar{b}(\bar{a}\bar{x}) g = g$. Setzen wir $\bar{b} g = \bar{b}(\bar{a}\bar{x}) h = r$, so folgt wegen der *Kommutativität* von $\overline{\mathfrak{X}}$

$$\bar{a}\bar{x}(h - r) = 0.$$

r und $f = h - r$ liefern eine Zerlegung der gewünschten Art.

Man kann im allgemeinen nicht damit rechnen, daß die Zerlegung (1) eindeutig bestimmt ist. Dies zeigt das

Beispiel 1. $\mathfrak{H} = \{[\xi, \eta]\}$ sei zweidimensional und $\mathfrak{G}$ bestehe aus den drei Transformationen e, x, y mit den Matrizendarstellungen

$$e = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \qquad x = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \qquad y = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}.$$

Dann ist $\mathfrak{R} = \{[\xi, \eta] | \eta = 0\}$

$$\mathfrak{F} = \{[\xi, \eta] | \xi = 0\} \cup \{[\xi, \eta] | \xi + \eta = 0\}.$$

Man erkennt sofort die im allgemeinen verschiedenen Zerlegungen

$$\begin{aligned} [\xi, \eta] &= [\xi, 0] + [0, \eta] \\ &= [\xi + \eta, 0] + [-\eta, \eta] \end{aligned}$$

vom Typus (1). Zugleich sieht man, daß $\mathfrak{F}$ kein linearer Raum ist. Ersetzt man die obigen Matrizen durch ihre Transponierten, so wird $\mathfrak{F}$, aber nicht $\mathfrak{R}$ ein linearer Raum.

Allgemein gilt der

Satz 1.2. *Es gilt* $\mathfrak{R} \cap \mathfrak{F} = \{0\}$.

Er ist unschwer zu beweisen. Unser nächstes Ziel ist der

Satz 1.3 (Aufspaltungssatz I). *Sind der reflexive Banach-Raum* $\mathfrak{H}$ *und sein dualer Raum* $\mathfrak{H}'$ *(je für sich) stark-konvex normierbar, derart, daß dann jeweils* $\Gamma = 1$ *gesetzt werden kann, so sind* $\mathfrak{R}$ *und* $\mathfrak{F}$ $\mathfrak{G}$*-invariante lineare abgeschlossene Teilräume von* $\mathfrak{H}$. *Jeder Vektor* $h \in \mathfrak{H}$ *besitzt genau eine Zerlegung*

$$h = r + f, \quad r \in \mathfrak{R},\ f \in \mathfrak{F},$$

wobei r *und* f *linear,* $\mathfrak{G}$*-invariant und norm-stetig von* h *abhängen. M.a.W.* $\mathfrak{H}$ *zerfällt direkt in* $\mathfrak{R}$ *und* $\mathfrak{F}$. — *Die von* $\overline{\mathfrak{G}}$ *in* $\mathfrak{R}$ *induzierte Halbgruppe von linearen Transformationen ist eine Gruppe.*

Hierbei heißt eine Norm $\|h\|$ in $\mathfrak{H}$ — wie üblich — *stark-konvex*, wenn aus

$$\|f\| \leq 1,\ \|g\| \leq 1,\ f \neq g$$

stets $\left\|\frac{f+g}{2}\right\| < 1$ folgt. Es gibt stark-konvex normierbare, nicht-reflexive Banach-Räume (DAY [4]).

§ 2. Beweis des Aufspaltungssatzes I

Wir legen die Voraussetzungen des Satzes 1.3 in diesem Paragraphen fest zugrunde. Es gilt dann der

Satz 2.1. *Ist* $\mathfrak{U}$ *eine Unterhalbgruppe von* $\mathfrak{G}$, *so ist jeder* $\mathfrak{G}$*-reversible Vektor auch* $\mathfrak{U}$*-reversibel.*

Beweis. Es sei $r \in \mathfrak{R}$. Jedenfalls gibt es eine Zerlegung

$$r = r_0 + f_0$$

von r in einen $\mathfrak{U}$-reversiblen Vektor r_0 und einen $\mathfrak{U}$-Fluchtvektor f_0 (Satz 1.1, auf $\mathfrak{U}$ angewendet). Es gibt also ein $\bar{u} \in \overline{\mathfrak{U}}$ mit $\bar{u} f_0 = 0$, d. h. $\bar{u} r = \bar{u} r_0 = \bar{u}\left(\frac{r + r_0}{2}\right)$. Wir können $\|r\| = 1$ annehmen und haben wegen $\Gamma = 1$ $\|r_0\| \leq 1$. Aus der Reversibilität und aus $\Gamma = 1$ folgt sogleich $\|r_0\| = \|r\| = \|\bar{u} r\| = \|\bar{u} r_0\|$ (vgl. JACOBS [12]). Als kann wegen $\Gamma = 1$ nur $\left\|\frac{r_0 + r}{2}\right\| = 1$ gelten. Aus der starken Konvexität der Norm folgt nun $r = r_0$, q. e. d.

Nun kann man die ganzen in [12] angestellten *Dualitätsbetrachtungen* auf unseren Fall übertragen. Wir nehmen die in [12] eingeführten Bezeichnungen und wiederholen alle bisherigen Überlegungen bezüglich der dualen Halbgruppe $\mathfrak{G}'$ im dualen Raum $\mathfrak{H}'$. Wir gewinnen eine Zerlegung

$$h' = r' + f', \quad r' \in \mathfrak{R}',\ f' \in \mathfrak{F}',$$

wobei $\mathfrak{R}'$ und $\mathfrak{F}'$ analog wie $\mathfrak{R}$ und $\mathfrak{F}$ definiert sind. Mittels Satz 2.1 gelingt nun der Beweis für den

Satz 2.2. *Ist* $r \in \mathfrak{R}$, $f' \in \mathfrak{F}'$, *so gilt* $(r, f') = 0$; *ist* $f \in \mathfrak{F}$, $r' \in \mathfrak{R}'$, *so gilt* $f, r') = 0$.

Beweis. Sei $r \in \mathfrak{R}$, $f' \in \mathfrak{F}'$ und $\bar{x}' \in \overline{\mathfrak{G}}'$ so bestimmt, daß $\bar{x}' f' = O' \in \mathfrak{H}'$ gilt. $\bar{x}'$ ist die duale Transformation einer Transformation $\bar{x} \in \overline{\mathfrak{G}}$. Es sei $\mathfrak{U} = \{\bar{x}^\nu | \nu = 0, 1, \ldots\}$. Dann ist r $\mathfrak{U}$-reversibel, es gibt also ein $\bar{y} \in \overline{\mathfrak{U}}$ mit $\bar{y}\bar{x}r = r$. $\overline{\mathfrak{U}}$ ist aber *abelsch*. Somit ist auch $\bar{x}\bar{y}r = r$. Wir haben damit

$$(r, f') = (\bar{x}\bar{y}r, f') = (\bar{y}r, \bar{x}' f') = 0.$$

Genau so beweist man die zweite Hälfte des Satzes.

Nun kann man den Aufspaltungssatz I nach genau demselben Verfahren beweisen, wie dies — unter etwas anderen Voraussetzungen — in [*12*] durchgeführt wurde ([*12*], Kap. III, § 4).

Man gewinnt außerdem aufs neue den

Satz 2.3 (Allgemeiner Ergodensatz von ALAOGLU-BIRKHOFF). *Für jedes* $h \in \mathfrak{H}$ *enthält die stark-abgeschlossene konvexe Hülle* $\mathfrak{K}(h)$ *der Menge* $\mathfrak{G}h = \{xh | x \in \mathfrak{G}\}$ *genau einen* $\mathfrak{G}$*-Fixpunkt* h_0, *d.h. genau einen Punkt* h_0 *mit*

$$x h_0 = h_0 \qquad \text{für alle } x \in \mathfrak{G}$$

(vgl. BIRKHOFF [2], ALAOGLU-BIRKHOFF [*1*]).

Beweis. Es sei

$$\mathfrak{K} = \left\{ \sum_{\nu=1}^{n} \alpha_\nu x_\nu \,|\, \alpha_\nu \geq 0,\ \sum_{\nu=1}^{n} \alpha_\nu = 1,\ x_\nu \in \mathfrak{G} \right\}$$

die *konvexe* Hülle der Halbgruppe $\mathfrak{G}$. $\mathfrak{K}$ ist eine Halbgruppe, deren Elemente bei einer geeigneten starkkonvexen Norm die gemeinsame Schranke 1 besitzen. Es gilt: Jeder $\mathfrak{K}$-reversible Vektor ist ein $\mathfrak{G}$-Fixpunkt (die Umkehrung dieser Aussage ist trivialerweise richtig). Denn ist r $\mathfrak{K}$-reversibel und $a \in \mathfrak{G}$ beliebig gewählt, ist ferner $\mathfrak{U}$ die konvexe Hülle der Halbgruppe $\mathfrak{A} = \{a^\nu | \nu = 0, 1, \ldots\}$, so ist r auch $\mathfrak{U}$-reversibel, denn $\mathfrak{U}$ ist eine Unterhalbgruppe von $\overline{\mathfrak{K}}$. $\overline{\mathfrak{U}}$ ist abelsch. Hieraus folgert man: Ist $\bar{b} \in \overline{\mathfrak{U}}$, so ist

$$\|ar - r\| \leq \|a\bar{b}r - \bar{b}r\|.$$

Denn $\|ar - r\| = \|a\bar{c}\bar{b}r - \bar{c}\bar{b}r\|$ für geeignetes $\bar{c} \in \overline{\mathfrak{U}}$. Es geht weiter

$$= \|\bar{c}(a\bar{b}r - \bar{b}r)\| \leq \|a\bar{b}r - \bar{b}r\|.$$

Setzt man speziell $\bar{b} = \frac{1}{n} \sum_{\nu=1}^{n} a^\nu$, so folgt

$$\|ar - r\| \leq \frac{2}{n} \|r\|.$$

Da n beliebig war, ist $\|ar - r\| = 0$, d.h. $ar = r$, was zu beweisen war.

§ 3. Formulierung des Aufspaltungssatzes II

Unsere eigentliche Absicht ist es, die Vektoren $h \in \mathfrak{H}$ in eine *fastperiodische* und eine Fluchtkomponente aufzuspalten. Der Begriff „fastperiodischer Vektor unter $\mathfrak{G}$" ist in [*12*] erklärt worden. Bezeichnen wir die Gesamtheit aller unter $\mathfrak{G}$ fastperiodischen Vektoren mit $\mathfrak{P}$, so ist $\mathfrak{P}$ ein $\mathfrak{G}$-invarianter abgeschlossener linearer Teilraum — nicht nur von $\mathfrak{H}$, sondern sogar von $\mathfrak{R}$ ([*12*], Satz II. 1.4). Wir formulieren den

Satz 3.1 (Aufspaltungssatz II). *Es sei $\mathfrak{H}$ ein reflexiver Banach-Raum; mindestens einer der beiden zueinander dualen Banach-Räume $\mathfrak{H}$, $\mathfrak{H}'$ sei gleichmäßig-konvex normierbar, derart, daß $\Gamma = 1$ gesetzt werden kann; der andere von den beiden Räumen soll unter der gleichen Nebenbedingung stark-konvex normierbar sein. Dann gilt $\mathfrak{R} = \mathfrak{P}$. Genauer: $\mathfrak{P}$ und $\mathfrak{F}$ sind $\mathfrak{G}$-invariante abgeschlossene lineare Teilräume von $\mathfrak{H}$; jeder Vektor $h \in \mathfrak{H}$ besitzt genau eine Zerlegung*

$$h = p + f, \quad p \in \mathfrak{P},\ f \in \mathfrak{F}.$$

Die Komponenten p, f hängen linear, $\mathfrak{G}$-invariant und normstetig von h ab. Die von der von $\mathfrak{G}$ in $\mathfrak{P}$ induzierten Halbgruppe gebildete starke Hülle (innerhalb von $\mathfrak{P}$) ist eine Gruppe.

Hierbei bezeichnen wir eine Norm $\|h\|$ in $\mathfrak{H}$ — wie üblich — als *gleichmäßig-konvex*, wenn es zu jedem $\varepsilon > 0$ ein $\delta(\varepsilon) > 0$ mit

$$\left\| \frac{f+g}{2} \right\| \leq 1 - \delta(\varepsilon) \quad \text{falls} \quad \|f\| \leq 1,\ \|g\| \leq 1,\ \|f-g\| \geq \varepsilon$$

gibt (vgl. DAY [*4*]); gleichmäßig-konvex normierbare Banach-Räume sind stets reflexiv (PETTIS [*18*]); jeder Hilbert-Raum ist gleichmäßig-konvex normiert (SZ.-NAGY [*17*]).

§ 4. Beweis des Aufspaltungssatzes II

Wir legen in diesem Paragraphen die Voraussetzungen des Satzes 3.1 (Aufspaltungssatz II) zugrunde. Unsere wesentlichen Beweismittel sind die folgenden beiden Sätze.

Satz 4.1. *Ist $\|h\|$ eine gleichmäßig-konvexe Norm in einem Banach-Raum $\mathfrak{B} = \{h, \ldots\}$, und ist $\mathfrak{M}$ eine schwach-abgeschlossene Menge konstanter Norm (d.h. gilt $\|g\| = \|h\|$ für $g, h \in \mathfrak{M}$), so ist $\mathfrak{M}$ totalbeschränkt (d.h. es gibt zu jedem $\varepsilon > 0$ endlichviele Punkte $h_1, \ldots, h_r \in \mathfrak{M}$ derart, daß zu jedem $h \in \mathfrak{M}$ ein ϱ mit $\|h - h_\varrho\| < \varepsilon$ existiert).*

Zum Beweis vgl. JACOBS [*11*], [*12*].

Satz 4.2. *Ist $\mathfrak{T}$ eine Gruppe von linearen Transformationen im Banach-Raum $\mathfrak{B} = \{h, \ldots\}$, welche eine gemeinsame Schranke besitzen, so ist ein Vektor $h \in \mathfrak{B}$ dann und nur dann fastperiodisch unter $\mathfrak{T}$, wenn die Menge $\mathfrak{T}h = \{th \mid t \in \mathfrak{T}\}$ totalbeschränkt (vgl. Satz 4.1) ist.*

Zum Beweis vgl. JACOBS [*12*].

Setzt man für $\mathfrak{B}$ den Raum der reversiblen Vektoren in demjenigen der beiden Banach-Räume $\mathfrak{H}$, $\mathfrak{H}'$, welcher gleichmäßig-konvex normierbar ist, so

induziert $\overline{\mathfrak{G}}$ bzw. $\overline{\mathfrak{G}}'$ in $\mathfrak{B}$ eine *Gruppe* $\mathfrak{T}$ von linearen Transformationen, die bei passender Wahl der Norm alle die Schranke 1 besitzen. Man sieht, daß dann für jedes $h \in \mathfrak{B}$ die Menge $\mathfrak{T}h$ schwach-abgeschlossen und von konstanter Norm ist; sie ist also totalbeschränkt (Satz 4.1), die Vektoren $h \in \mathfrak{B}$ sind alle *fastperiodisch* unter $\mathfrak{T}$ (Satz 4.2). Nun kann man aus der Definition des Begriffs „fastperiodischer Vektor unter einer Halbgruppe" [*12*] leicht entnehmen, daß die Vektoren $h \in \mathfrak{B}$ auch unter jeder Unterhalbgruppe von $\mathfrak{B}$ fastperiodisch sind, insbesondere unter der von $\mathfrak{G}$ (bzw. $\mathfrak{G}'$) in $\mathfrak{B}$ induzierten Halbgruppe. Von hier aus schließt man leicht auf die Fastperiodizität des Vektoren $h \in \mathfrak{B}$ unter der Halbgruppe $\mathfrak{G}$ (bzw. $\mathfrak{G}'$).

Wir haben somit gefunden, daß in mindestens einem der beiden Räume $\mathfrak{H}$, $\mathfrak{H}'$ die Begriffe „reversibel" und „fastperiodisch" sich decken. Ist $\mathfrak{H}$ dieser Raum, so sind wir fertig. Ist $\mathfrak{H}'$ dieser Raum, so haben wir noch nachzuweisen, daß aus $\mathfrak{P}' = \mathfrak{R}'$ sogleich $\mathfrak{P} = \mathfrak{R}$ folgt. Nach den voranstehenden Betrachtungen genügt es offenbar, anzunehmen, daß $\mathfrak{H} = \mathfrak{R}$ und $\overline{\mathfrak{G}}$ eine Gruppe ist. Wir bemerken zunächst, daß für und beliebiges

$$h' = p' + f', \quad p' \in \mathfrak{P}', \; f' \in \mathfrak{F}', \; \bar{x} \in \overline{\mathfrak{G}}$$
$$(\bar{x}\, r, h') = (\bar{x}\, r, p') = (r, \bar{x}'\, p')$$

gilt. Aus der $\overline{\mathfrak{G}}'$-Fastperiodizität von p' folgt nun leicht die Fastperiodizität der skalaren Funktion $\varphi(\bar{x}) = (\bar{x}\, r, h')$ auf der Gruppe $\overline{\mathfrak{G}}$. Dies gilt für beliebige $h' \in \mathfrak{H}'$. Wir sind somit fertig, sobald der folgende Satz bewiesen ist:

Satz 4.3. *Ein Vektor $h \in \mathfrak{H}$ ist dann und nur dann fastperiodisch unter der Gruppe $\overline{\mathfrak{G}}$, wenn für jedes $h' \in \mathfrak{H}'$ die (skalare) Funktion $\varphi(\bar{x}) = (\bar{x}\, h, h')$ fastperiodisch auf der Gruppe $\mathfrak{G}$ ist.*

Ein Beweis dieses Satzes für *abelsche* Gruppen $\overline{\mathfrak{G}}$ findet sich in [*12*]; die Übertragung auf den allgemeinen Fall bereitet keinerlei Schwierigkeiten.

Es bleibt nun noch zu zeigen, daß die von $\mathfrak{G}$ in $\mathfrak{P}$ induzierte Halbgruppe eine *Gruppe* zur *starken* Hülle hat. Dies wurde in [*12*] bewiesen.

Damit ist der Aufspaltungssatz II vollständig bewiesen. Außer der Aufspaltung von $\mathfrak{H}$ erhält man eine analoge Aufspaltung von $\mathfrak{H}'$. Zwischen den beiden Aufspaltungen bestehen *Orthogonalitätsrelationen* (vgl. Satz 2.2; [*12*]).

§ 5. Kontraktionen im Hilbert-Raum

Die in [*12*], Anhang III durchgeführten Betrachtungen sind von irgendwelchen Kommutativitätseigenschaften von $\mathfrak{G}$ unabhängig und lassen sich somit auf unseren Fall übertragen. Wir gewinnen den

Satz 5.1. *Ist $\mathfrak{H}$ ein Hilbert-Raum und $\mathfrak{G}$ eine Halbgruppe von nichtdehnenden Transformationen in $\mathfrak{H}$, so ist $\mathfrak{P}' = \mathfrak{P}$, $\mathfrak{F}' = \mathfrak{F}$, und $\mathfrak{P}$ und $\mathfrak{F}$ sind orthogonalkomplementär in $\mathfrak{H}$. $\mathfrak{G}$ induziert in $\mathfrak{P}$ eine unitäre Halbgruppe $\mathfrak{G}_0$. Desgleichen induziert $\mathfrak{G}'$ in $\mathfrak{P}$ eine unitäre Halbgruppe $\mathfrak{G}_0'$. Die starken Hüllen von $\mathfrak{G}_0$ bzw. $\mathfrak{G}_0$ in $\mathfrak{P}$ stimmen überein; es handelt sich um eine unitäre Gruppe in $\mathfrak{P}$.*

§ 6. Der Ergodizitätscharakter nichtkommutativer Halbgruppen in Banach-Räumen

Sowohl aus den durch den Satz 3.1 (Aufspaltungssatz II) festgestellten Fastperiodizitätseigenschaften von $\mathfrak{G}$ in $\mathfrak{H}$ als auch durch eine direkte Schlußweise (vgl. Satz 2.3) kann man den Ergodensatz gewinnen. Danach ist jedem $h \in \mathfrak{H}$ genau ein $\mathfrak{G}$-Fixpunkt h_0 als *Mittelwert* zugeordnet. Es erhebt sich die Frage, in welcher Weise h_0 durch *Näherungsmittel* $\frac{1}{n}\sum_{\nu=1}^{n} x_\nu h$ $(x_\nu \in \mathfrak{G})$ approximiert werden kann.

Ist $\mathfrak{G}$ *abelsch*, so ist die Antwort einfach: Für jedes $h \in \mathfrak{H}$ ist die auf $\mathfrak{G}$ definierte Funktion $h(x) = xh$ (mit Werten aus $\mathfrak{H}$) *ergodisch*, d.h. es gibt zu jedem $\varepsilon > 0$ endlichviele Elemente $x_1, \ldots, x_n$, derart, daß

$$\left\| \frac{1}{n}\sum_{\nu=1}^{n} h(x x_\nu) - h_0 \right\| < \varepsilon \qquad (x \in \mathfrak{G})$$

gilt. Hat $\mathfrak{G}$ z.B. die Gestalt $\mathfrak{G} = \{a^\nu \mid \nu = 0, 1, \ldots\}$, so kann man verschärfend schließen, daß

$$\lim_{n\to\infty} \left\| \frac{1}{n}\sum_{\nu=1}^{n} x a^\nu h - h_0 \right\| = 0 \qquad \text{(gleichmäßig in } x \in \mathfrak{G})$$

gilt.

Ist $\mathfrak{G}$ *nicht-abelsch*, so ergeben sich verschiedene Möglichkeiten, den oben formulierten Begriff „ergodisch" zu verallgemeinern. Die Funktion $h(x)$ heißt *linksergodisch* zum Mittelwert h_0, wenn es zu jedem $\varepsilon > 0$ Elemente $x_1, \ldots, x_n \in \mathfrak{G}$ mit

$$\left\| \frac{1}{n}\sum_{\nu=1}^{n} h(x x_\nu) - h_0 \right\| < \varepsilon \qquad (x \in \mathfrak{G})$$

gibt. Kann man stets

$$\left\| \frac{1}{n}\sum_{\nu=1}^{n} h(x_\nu x) - h_0 \right\| < \varepsilon \qquad (x \in \mathfrak{G})$$

erreichen, so heißt die Funktion $h(x)$ *rechts-ergodisch*. Ist die Funktion $h(x)$ sowohl links- als auch rechts-ergodisch, so heißt sie schlechthin *ergodisch*. Kann man stets

$$\left\| \frac{1}{n}\sum_{\nu=1}^{n} h(x x_\nu y) - h_0 \right\| < \varepsilon \qquad (x, y \in \mathfrak{G})$$

erreichen, so heißt $h(x)$ *voll-ergodisch* (oder auch stark-ergodisch, vgl. GODEMENT [10], MAAK [16], KEINER [13]).

Satz 2.3 besagt, daß unter den angegebenen Voraussetzungen $h(x)$ *stets links-ergodisch* ist. Ist h $\mathfrak{G}$-fastperiodisch, so ist $h(x)$ *voll-ergodisch*. Nach dem Aufspaltungssatz II könnte man somit hoffen, daß $h(x)$ stets voll-

ergodisch ist. Jedoch muß man im allgemeinen damit rechnen, daß die Fluchtkomponenten *nicht* einmal *rechts-ergodisch* sind. Dies zeigt das

Beispiel 2. $\mathfrak{H}$ sei ein Hilbert-Raum von abzählbarer Dimension, $e_1, e_2, \ldots$ eine orthonormierte Basis von $\mathfrak{H}$. Ist τ eine Permutation der Menge aller natürlichen Zahlen $\nu = 1, 2, \ldots$, so gibt es genau eine unitäre Transformation t von $\mathfrak{H}$ mit

$$t e_\nu = e_{\tau(\nu)} \qquad (\nu = 1, 2, \ldots).$$

$\mathfrak{G}$ sei nun die Gruppe all derjenigen t, die zu Permutationen τ gehören, welche alle ν bis auf endlichviele festlassen. Dann sind alle Voraussetzungen des Aufspaltungssatzes II erfüllt. Man sieht sofort, daß $\mathfrak{P} = \{0\}$, $\mathfrak{H} = \mathfrak{F}$ gilt (vgl. JACOBS [12]). Man folgert sogleich, daß $h_0 = 0$ für alle $h \in \mathfrak{H}$ gilt. Wir zeigen, daß z.B. $e_1(x)$ nicht rechts-ergodisch ist. Es seien $x_1, \ldots, x_n \in \mathfrak{G}$ beliebig gewählt. Dann gibt es ein μ mit $x_\nu e_\mu = e_\mu$ $(\nu = 1, \ldots, n)$, sowie ein $x \in \mathfrak{G}$ mit $x e_1 = e_\mu$. Man sieht sogleich, daß

$$\frac{1}{n} \sum_{\nu=1}^{n} x_\nu x e_1 = e_\mu$$

gilt, woraus die Behauptung unmittelbar folgt. Analog sieht man ein, daß $h(x)$ für $h \neq 0$ niemals rechts-ergodisch ist.

Man kann immerhin zeigen, daß die $h \in \mathfrak{H}$ mit voll-ergodischem $h(x)$ einen abgeschlossenen, $\mathfrak{G}$-invarianten linearen Teilraum von $\mathfrak{H}$ bilden; man könnte versuchen, einen komplementären Teilraum zu finden, und so in Form eines neuen Aufspaltungssatzes weitere Aufklärung über das Ergodozitätsverhalten der $h(x)$ im Fall nicht-kommutativer Halbgruppen $\mathfrak{G}$ zu gewinnen; hierüber ist mir nichts Näheres bekannt.

Verändert man die eben definierten Ergodizitätsbegriffe in der Weise, daß man nicht die *starke* (Norm-)Topologie, sondern die *schwache* Topologie zugrunde legt („schwache" Ergodizitätsbegriffe), so kann man leicht zeigen, daß $h(x)$ stets *schwach-ergodisch* ist. Aus Beispiel 2 ist jedoch leicht zu ersehen, daß $h(x)$ im allgemeinen *nicht schwach-vollergodisch* ist.

Literatur

[1] ALAOGLU, L., and G. BIRKHOFF: General ergodic theorems. Ann. of Math. **41**, 293—309 (1940). — [2] BIRKHOFF, G.: An ergodic theorem for general semigroups. Proc. Nat. Acad. Sci. U.S.A. **25**, 625—627 (1939). — [3] BOURBAKI, N.: Espaces vectoriels topologiques. Paris 1953. — [4] DAY, M. M.: Reflexive spaces not isomorphic to uniformly convex spaces. Bull. Amer. Math. Soc. **47**, 313—317 (1941). — [5] DAY, M. M.: Ergodic theorems for abelian semigroups. Trans. Amer. Math. Soc. **51**, 399—412 (1942). — [6] DAY, M. M.: Uniform convexity in factor and conjugate spaces. Ann. of Math. **45**, 375—385 (1944). — [7] DAY, M. M.: Means for the bounded functions and ergodicity of the bounded representations of semi-groups. Trans. Amer. Math. Soc. **69**, 276—291 (1950). — [8] DUNFORD, N., and S. MILLER: On the ergodic theorem. Trans. Amer. Math. Soc. **60**, 538—549 (1946). — [9] EBERLEIN, W. F.: Abstract ergodic theorems and weakly almost periodic functions. Trans. Amer. Math. Soc. **67**, 217—250 (1949). — [10] GODEMENT, R.: Les fonctions de type positif et la théorie des groupes. Trans. Amer. Math.

Soc. **63**, 1—84 (1948). — [*11*] JACOBS, K.: Periodizitätseigenschaften beschränkter Gruppen im Hilbertschen Raum. Math. Z. **61**, 408—428 (1955). — [*12*] JACOBS, K.: Ergodentheorie und fastperiodische Funktionen auf Halbgruppen. Math. Z. **64**, 298—338 (1956). — [*13*] KEINER, H.: Verallgemeinerte fastperiodische Funktionen anf Halbgruppen. Arch. Math. 1957. — [*14*] LORCH, E. R.: Means of iterated transformations in reflexive vector spaces. Bull. Amer. Math. Soc. **45**, 945—947 (1939). — [*15*] MAAK, W.: Fastperiodische Funktionen auf Halbgruppen. Acta mathematica **87**, 33—57 (1952). — [*16*] MAAK, W.: Integralmittelwerte von Funktionen auf Gruppen und Halbgruppen. J. reine u. angew. Math. **190**, 34—48 (1952). — [*17*] SZ.-NAGY, B.: Spektraldarstellungen linearer Transformationen des Hilbertschen Raumes. Berlin 1942. — [*18*] PETTIS: A proof that every uniformly convex space is reflexive. Duke Math. J. **5**, 249—253 (1939). — [*19*] RIESZ, F.: Some mean ergodic theorems. J. London Math. Soc. **13**, 274—278 (1938). — [*20*] RIESZ, F.: Sur la théorie ergodique des espaces abstraits. Acta Sci. Math. Szeged **10**, 1—20 (1941). — [*21*] RIESZ, F.: Another proof of the mean ergodic theorem. Acta Sci. Math. Szeged **10**, 75—76 (1941). — [*22*] RIESZ, F., u. B. SZ.-NAGY: Leçons d'analyse fonctionnelle. Budapest 1952. — [*23*] YOSIDA, K., and S. KAKUTANI: Operator-theoretical treatment of Markoff process and mean ergodic theorem. Ann. of Math. **42**, 188—228 (1941). — [*24*] DIXMIER, J.: Les moyennes invariantes dans les semi-groupes et leurs applications. Acta Sci. Math. Szeged **12**, 213—227 (1950).

Mathematisches Institut der Universität München

(Eingegangen am 9. September 1956)

BAUER, F. L. u. K. SAMELSON
Math. Zeitschr. Bd. 67, S. 93—98 (1957)

Polynomkerne und Iterationsverfahren

Herrn GEORG FABER zum 80. Geburtstag am 5. April 1957 gewidmet

Von

FRIEDRICH L. BAUER und **KLAUS SAMELSON**

Die ursprüngliche BERNOULLIsche Methode hat in jüngster Zeit beträchtliche Ausweitungen erfahren. Die BERNOULLIsche Verfahrensklasse umfaßt heute einen breiteren Anwendungsbereich (z.B. Faktorisierung von Polynomen, simultane Nullstellenbestimmung, Bestimmung von Matrixeigenwerten und -Vektoren) und mannigfache Varianten der numerischen Durchführung, wie z.B. den *QD*-Algorithmus [4] und die *LR*-Transformation [5] von RUTISHAUSER oder die Treppeniteration des einen der Verf. [2], [3]. Unter den zahlreichen Querverbindungen zu algebraischen Problemkreisen sei ein Zusammenhang der Treppeniteration mit der allgemeinen Theorie der Orthogonalpolynome herausgegriffen, insbesondere eine Beziehung zu der Formel von CHRISTOFFEL-DARBOUX für Polynomkerne[1]). Als heuristisches Mittel führt sie auf eine Konvergenzbeschleunigung der Treppeniteration; es ergibt sich im Limes eine Konvergenz wie die des NEWTON-Verfahrens (bei gleichem Arbeitsaufwand). Die offensichtlichen numerischen Vorzüge der Treppeniteration werden jedoch aufrecht erhalten.

1. Variante der BERNOULLI-Iteration

Wir gehen von einer Variante [1] der klassischen BERNOULLIschen Methode aus, die zwanglos auf das Verfahren der Treppeniteration [2] führt. Dabei wählen wir eine Schreibweise, die gegenüber unseren früheren Darstellungen nur unwesentlich modifiziert ist dadurch, daß $P(x)$ durch $x^n P\left(\frac{1}{x}\right)$ ersetzt wird. Hier, wie im folgenden, seien λ_μ, $\mu = 1 \dots n$, die Nullstellen von $P(x)$; es sei

$$|\lambda_n| \leq |\lambda_{n-1}| \leq \cdots \leq |\lambda_1|.$$

Rechenvorschrift I (BERNOULLI-LAGRANGE-*Variante*):

Gegeben: $P(x) \in \overline{\pi}_n$ [2])

Ausgangsnäherung: $q^{(0)}(x) \in \pi_{n-1}$

Algorithmus: Für $i = 0, 1, 2, \dots$

$$\frac{P(x) - \dfrac{P(0)}{q^{(i)}(0)}\, q^{(i)}(x)}{x} = q^{(i+1)}(x).$$

[1]) Siehe etwa G. SZEGÖ, Orthogonal Polynomials. Amer. Math. Soc. Colloquium Publications. Band 23 (1939).

[2]) π_n sei der Raum der Polynome vom Höchstgrad n, $\overline{\pi}_n$ der Raum der Polynome vom genauen Grad n, deren führender Koeffizient 1 ist.

Satz I. *Im Algorithmus der Rechenvorschrift I streben mit linearer Konvergenz*

$$\left.\begin{aligned} -\frac{P(0)}{q^{(i)}(0)} &\to \lambda_n \\ q^{(i)}(x) &\to \frac{P(x)}{x-\lambda_n} \in \bar{\pi}_{n-1} \end{aligned}\right\} \text{ für } i \to \infty$$

falls $|\lambda_n| < |\lambda_\mu|$, $\mu \neq n$ *und* $q^{(0)}(\lambda_n) \neq 0$.

Falls $\lambda_n = 0$, ist der Algorithmus stationär und konvergiert trivialerweise. Andernfalls gilt in der Umgebung von $z = 0$ die TAYLOR-Entwicklung

$$R(z) = \frac{q^{(0)}(z)}{P(z)} = \frac{q^{(0)}(0)}{P(0)} - \frac{q^{(0)}(0)\, q^{(1)}(0)}{(P(0))^2} z + \frac{q^{(0)}(0)\, q^{(1)}(0)\, q^{(2)}(0)}{(P(0))^3} z^2 - + \cdots,$$

aus der, durch Vergleich mit der nach geometrischen Reihen entwickelten und aufsummierten Partialbruchzerlegung von $R(z)$, die Aussagen von Satz I und der Zusammenhang mit dem klassischen BERNOULLI-Verfahren sofort ersichtlich sind. Es ergibt sich auch, daß der Algorithmus nur versagt, wenn in der obenstehenden TAYLOR-Reihe ein Koeffizient verschwindet.

2. Treppeniteration

Ein zweiter Lauf der BERNOULLI-Iteration, mit $\frac{P(x)}{x-\lambda_n}$ anstatt $P(x)$ liefert $\frac{P(x)}{(x-\lambda_n)(x-\lambda_{n-1})}$, falls $|\lambda_{n-1}| < |\lambda_\mu|$, $\mu \neq n$, $n-1$. Ein dritter, vierter, ... l-ter Lauf ($l \leq n$) kann sich entsprechend anschließen.

Die Treppeniteration unterscheidet sich von diesem Vorgehen nur insofern, als bereits die jeweiligen Zwischenergebnisse des k-ten Iterationslaufs an Stelle von $\frac{P(x)}{(x-\lambda_n)\cdots(x-\lambda_{n-k+1})}$ zur Berechnung der $(k+1)$-ten Iterationsreihe dienen.

Wir gelangen so zur

Rechenvorschrift II (Treppeniteration):

Gegeben: $q_n^{(i)}(x) \equiv P(x) \in \bar{\pi}_n$

Ausgangsnäherungen: $q_\mu^{(0)}(x) \in \pi_\mu$, $\mu = n-1, n-2, \ldots, n-l \geq 0$

Algorithmus: Für $i = 0, 1, 2, \ldots$

für $\mu = n-1, n-2, \ldots, n-l$

$$\frac{q_{\mu+1}^{(i)}(x) - \frac{q_{\mu+1}^{(i)}(0)}{q_\mu^{(i)}(0)} q_\mu^{(i)}(x)}{x} = q_\mu^{(i+1)}(x).$$

Satz II. *Im Algorithmus der Rechenvorschrift II strebt, abgesehen von gewissen singulären Fällen*[3] *für* $i \to \infty$ $(\mu = n-1, n-2, \ldots, n-l)$

$$-\frac{q_{\mu+1}^{(i)}(0)}{q_{\mu}^{(i)}(0)} \to \lambda_{\mu+1} \quad \textit{falls} \quad |\lambda_{\mu+2}| < |\lambda_{\mu+1}| < |\lambda_{\mu}|\ [4])$$

$$q_{\mu}^{(i)}(x) \to (x-\lambda_1)(x-\lambda_2)\ldots(x-\lambda_{\mu}) \in \bar{\pi}_{\mu} \quad \textit{falls} \quad |\lambda_{\mu+1}| < |\lambda_{\mu}|\ [5]).$$

Man beachte, daß die Reihe der $q_k^{(i)}(x)$ konvergieren kann, ohne daß die vorhergehende Reihe es tut.

Der Konvergenzbeweis für diese simultane Iteration ist komplizierter als in 1.; vgl. [2], wo ebenfalls $P(x)$ durch $x^n P\left(\frac{1}{x}\right)$ zu ersetzen ist.

3. Polynomkerne

Die (reellen) Polynome $p_\nu(x) \in \bar{\pi}_\nu$, $\nu = 0, 1, 2, \ldots$, seien orthogonal zur Verteilung $d\alpha(x)$

$$\int p_\varrho(x)\, p_\sigma(x)\, d\alpha(x) = \delta_{\varrho\sigma} \|p_\varrho\|^2.$$

Als Polynomkerne bezeichnet man die Ausdrücke

$$K_\mu(x, y) = \sum_{\nu=0}^{\mu} \frac{p_\nu(x)\, p_\nu(y)}{\|p_\nu\|^2}$$

vermöge ihrer Eigenschaft, Einheitskerne im Raum π_μ zu sein:

$$\int K_\mu(x, y)\, g(y)\, d\alpha(y) = g(x) \quad \text{für} \quad g(x) \in \pi_\mu.$$

Es gilt bekanntlich die Formel von CHRISTOFFEL-DARBOUX

$$K_\mu(x, y) = \frac{1}{\|p_\mu\|^2} \frac{p_{\mu+1}(x)\, p_\mu(y) - p_\mu(x)\, p_{\mu+1}(y)}{x - y},$$

und die $K_\mu(x, y)$ sind bezüglich x wieder Orthogonalpolynome, jetzt aber zur Verteilung $(x-y)\,d\alpha(x)$.

Die Polynome $q_\mu^{(i)}(x)$, $\mu = n, n-1, \ldots, n-l$, seien aus einem Orthogonalsystem zur Verteilung $d\alpha_i(x)$. Die zugehörigen Polynomkerne lauten bis auf einen belanglosen Faktor

$$\frac{q_{\mu+1}^{(i)}(x)\, q_\mu^{(i)}(y^{(i)}) - q_{\mu+1}^{(i)}(y^{(i)})\, q_\mu^{(i)}(x)}{x - y^{(i)}}, \quad \mu = n-1, \ldots, n-l.$$

Vergleich mit Rechenvorschrift II ergibt, daß die durch einen Schritt der Treppeniteration sich ergebenden Polynome $q_\mu^{(i+1)}(x)$, $\mu = n-1, \ldots, n-l$ Polynomkerne mit $y^{(i)} \equiv 0$ sind. Sie sind damit Orthogonalpolynome zur Verteilung $x\, d\alpha_i(x) = x^{i+1} d\alpha_0(x)$. Auch $q_n^{(i+1)}(x) = P(x)$ gehört zu diesem Orthogonalsystem, wenn $P(x)$ nur einfache Nullstellen hat und die Belegung

[3]) Gekennzeichnet durch das Verschwinden eines der Ausdrücke

$M_\mu = \det|q_i^0(\lambda_{k+1})|$; $\mu = n-1, n-2, \ldots, n-k$; $i, k = \mu, \mu+1, \ldots, n-1$.

[4]) Die erste bzw. die zweite Ungleichung entfällt für $\mu = n-1$ bzw. $\mu = 0$.

[5]) Die Ungleichung entfällt für $\mu = 0$.

$\alpha_0(x) = \sum_{\mu=1}^{n} e_\mu Y(x-\lambda_\mu)$ eine Treppenfunktion ist mit Sprungstellen an den Nullstellen von $P(x)$, so daß also $d\alpha_i(x) = \sum_{\mu=1}^{n} e_\mu \lambda_\mu^i \cdot dY(x-\lambda_\mu)$ ist. All diese Orthogonalsysteme brechen mit $P(x)$ ab wegen $\|P(x)\| = 0$.

Falls $\alpha_0(x)$ monoton ist, also alle e_μ (reell und) positiv und ferner die Nullstellen von $P(x)$ positiv sind, ist das Grenzverhalten der Treppeniteration aus der klassischen Theorie der Orthogonalpolynome heraus leicht zu verstehen: $q_\mu^{(i)}(x)$ ist dann dasjenige Polynom aus $\bar{\pi}_\mu$, dessen Norm bezüglich des (monotonen) $\alpha_i(x)$ minimal ist. Mit wachsendem i erdrücken die zu den μ betragsgrößten Sprungstellen λ_ν gehörigen Gewichte $\lambda_\nu^i e_\nu$ alle anderen Gewichte, so daß die Minimallösung gegen ein Polynom mit eben diesen Sprungstellen als Nullstellen strebt.

Die getroffenen Voraussetzungen sind aber in Satz II ebensowenig notwendig wie gar die Voraussetzung, daß die Ausgangsnäherung ein Orthogonalsystem ist, also einer dreigliedrigen Rekursionsformel genügt.

4. NEWTON-artige Konvergenz

Nach diesen Vorbereitungen kommen wir zum Ziel unserer Arbeit. Nach dem vorigen Abschnitt liegt es zunächst nahe, bei der BERNOULLI-LAGRANGE-Variante von 1. die Auszeichnung $y_i = 0$ in der CHRISTOFFEL-DARBOUX-Formel fallen zu lassen und y_i jeweils gleich der besten schon erhaltenen Approximation für $\lambda_n (y_i = \beta_i)$ zu setzen. Man erhält damit

Rechenvorschrift III (Polynom-Iteration mit NEWTON*-artiger Konvergenz):*

Gegeben: $P(x) \in \bar{\pi}_n$

Ausgangsnäherungen: $q^{(0)}(x) \in \pi_{n-1}$; $\beta^{(0)}$

Algorithmus: Für $i = 0, 1, 2, \ldots$,

$$\frac{P(x) - \dfrac{P(\beta^{(i)})}{q^{(i)}(\beta^{(i)})} q^{(i)}(x)}{x - \beta^{(i)}} = q^{(i+1)}(x)$$

$$\beta^{(i)} - \frac{P(\beta^{(i)})}{q^{(i)}(\beta^{(i)})} = \beta^{(i+1)}.$$

Falls die Iteration nun noch konvergiert, strebt

$$q^{(i)}(x) \to \frac{P(x)}{x - \lambda_\nu}$$

und

$$\beta^{(i)} \to \lambda_\nu.$$

Ähnlich dem NEWTON-Verfahren herrscht jedoch nur noch dann Konvergenz, wenn die Ausgangsnäherungen $q^{(0)}(x)$ und β_0 in einer gewissen Umgebung einer Lösung liegen. Dafür kann jedoch nunmehr jede einfache Nullstelle quadratisch approximiert werden. Der Beweis kann nach klassischen

Methoden[6]) geführt werden, die elementare, aber etwas umständliche Durchführung der (mehrdimensionalen) TAYLOR-Entwicklung soll hier nicht wiedergegeben werden.

5. Treppeniteration mit NEWTON-artiger Konvergenz

Als entsprechende Verallgemeinerung der Treppeniteration wird man erwarten, daß nunmehr Polynomkerne mit den nach Rechenvorschrift III definierten $\beta^{(i)}$ die Iteration tragen. Dabei käme jedoch nur für die betragskleinste Nullstelle quadratische Konvergenz zustande. Wir lassen diesen Fall beiseite[7]) und wenden uns sogleich einer allgemeineren Rechenvorschrift zu, wobei $\beta^{(i)}$ nicht mehr von μ unabhängig ist, sondern [für die Bestimmung von $q_\mu^{(i+1)}(x)$] sich gerade aus der erzielten Approximation für $\lambda_{\mu+1}$ ergibt. Dabei geht auch der Zusammenhang mit Polynomkernen wieder gänzlich verloren. Wir haben die

Rechenvorschrift IV (Treppeniteration mit NEWTON*-artiger Konvergenz):*

Gegeben: $q_n^{(i)}(x) \equiv P(x) \in \bar{\pi}_n$

Ausgangsnäherungen: $\left.\begin{array}{l} q^{(0)}(x) \in \pi_\mu \\ \beta_\mu^{(0)} \end{array}\right\} \quad \mu = n-1, n-2, \ldots, n-l \geqq 0$

Algorithmus: Für $i = 0, 1, 2, \ldots$

für $\mu = n-1, n-2, \ldots, n-l$

$$\frac{q_{\mu+1}^{(i)}(x) - \dfrac{q_{\mu+1}^{(i)}(\beta_{\mu+1}^{(i)})}{q_\mu^{(i)}(\beta_{\mu+1}^{(i)})}\, q_\mu^{(i)}(x)}{x - \beta_{\mu+1}^{(i)}} = q_\mu^{(i+1)}(x)$$

$$\beta_{\mu+1}^{(i)} - \frac{q_{\mu+1}^{(i)}(\beta_{\mu+1}^{(i)})}{q_\mu^{(i)}(\beta_{\mu+1}^{(i)})} = \beta_{\mu+1}^{(i+1)}.$$

Im Konvergenzfall strebt

$$q_\mu^{(i)}(x) \to (x - \lambda_{\alpha_1})(x - \lambda_{\alpha_2}) \ldots (x - \lambda_{\alpha_\mu})$$

und

$$\beta_{\mu+1}^{(i)} \to \lambda_{\alpha_\mu}.$$

Konvergenz herrscht wiederum, wenn die Ausgangsnäherung in einer gewissen Umgebung einer Lösung ist, die Reihenfolge der Nullstellen braucht wiederum nicht mehr eine solche nach Beträgen sein. Einfache Nullstellen werden quadratisch approximiert, für den Beweis gilt die Bemerkung am Schluß von 4.

6. Diskussion der numerischen Durchführung

Gegenüber der Treppeniteration hat die Rechenvorschrift IV den Vorteil der im Limes quadratischen und damit erheblich rascheren Konvergenz. Sie ist also von großem Nutzen, wenn die linear konvergente Treppeniteration

[6]) Siehe etwa D. R. HARTREE, Numerical Analysis, Oxford 1952, p. 191.

[7]) Er entspricht der Methode der Nullpunktverschiebung beim QD-Algorithmus von RUTISHAUSER [3].

langsam konvergiert. Je schlechter jedoch die Ausgangsnäherung, um so weniger kann man sich praktisch auf die nur im Limes definierte quadratische Konvergenz verlassen. Es wird sich im allgemeinen empfehlen, durch einige wenige Schritte der gewöhnlichen Treppeniteration eine gewisse definierte Ausgangsnäherung herzustellen und dann mit quadratischer Konvergenz weiter zu iterieren. Geschieht dies nicht zu früh, so stellen sich die Nullstellen auch nach einer Ordnung der Beträge ein. Insbesondere ist stets zu beachten, daß nicht mehr für jede Ausgangsnäherung Konvergenz herrscht; es dürfte sich also schon deshalb empfehlen, nicht zu früh mit der Konvergenzbeschleunigung zu beginnen.

Literatur

[1] BAUER, F. L.: Quadratisch konvergente Durchführung der Bernoulli-Jacobischen Methode zur Nullstellenbestimmung. Sitzgsber. Bayer. Akad. Wiss. **1954**, 275—303. — [2] BAUER, F. L.: Direkte Faktorisierung eines Polynoms. Sitzgsber. Bayer. Akad. Wiss. **1956**, 163—203. — [3] BAUER, F. L.: Das Verfahren der Treppeniteration und verwandte Verfahren zur Lösung algebraischer Eigenwertprobleme. ZAMP (im Druck). — [4] RUTISHAUSER, H.: Der Quotienten-Differenzen-Algorithmus. Basel u. Stuttgart: Birkhäuser 1956. — [5] RUTISHAUSER, H.: Report on the Solution of Eigenvalue Problems with the *LR*-Transformation. In: NBS Appl. Math. Ser. Nr. 49, Washington (im Erscheinen).

München, Mathematisches Institut der Technischen Hochschule

(Eingegangen am 10. November 1956)

BIEBERBACH, L.
Math. Zeitschr. Bd. 67, S. 99—102 (1957)

Eine Bemerkung zur konformen Abbildung zweifach zusammenhängender Gebiete

Herrn GEORG FABER zum 80. Geburtstag am 5. April 1957 gewidmet

Von

LUDWIG BIEBERBACH

Zu den vielen methodisch wertvollen Gedanken, mit denen G. FABER die Mathematik bereichert hat, gehört sein Verfahren zur Untersuchung des Randverhaltens der konformen Kreisabbildung [*1*]. Daraus entwickelte sich die GRÖTZSche Flächenstreifenmethode und aus dieser wieder ist der Begriff der extremalen Länge entstanden. Er wird in dieser kleinen Note Verwendung finden.

Vor geraumer Zeit hat der Verf. [*2*] mit RIEMANNschen Methoden den RIEMANNschen Satz bewiesen, daß jedes n-fach zusammenhängende Gebiet umkehrbar eindeutig und konform auf eine n-blättrige Kreisscheibe abgebildet werden kann. H. GRUNSKY [*3*] hat einen mit einfacheren Hilfsmitteln auskommenden Beweis angegeben und auch die Unität geklärt.

Man kann sich fragen, welche Bedeutung die kreisförmige Begrenzung, und zwar durch den gleichen Kreis in allen Etagen bei dem RIEMANNschen Satz spielt. Die Antwort lautet: so gut wie keine, das Wesentliche ist die Verteilung der Verzweigungspunkte. Das will ich in dieser Note für den Fall zweifach zusammenhängender Gebiete begründen. Ich beweise nämlich:

Es seien C_1 und C_2 die Ränder zweier einfach zusammenhängender schlichter Gebiete G_1 und G_2. Der Durchschnitt $G_1 \cap G_2$ soll nicht leer sein. C_1 und C_2 sollen beide aus je mehr als einem Punkt bestehen und sollen mindestens einen gemeinsamen Punkt haben, der aus $G_1 \cap G_2$ längs einem gemeinsamen Weg erreichbar ist. Dann kann jedes vorgegebene zweifach zusammenhängende Gebiet ohne punktförmige Randkomponente umkehrbar eindeutig und konform auf ein zweiblättriges RIEMANN*sches Flächenstück G abgebildet werden, das im einen Blatt von einer über C_1 gelegenen Randkomponente, im anderen Blatt von einer über C_2 gelegenen Randkomponente begrenzt wird. Es entsteht, indem man G_1 und G_2 durch zwei ihrem Durchschnitt angehörige passend gewählte Verzweigungspunkte vereinigt*[1]).

Der Satz erinnert in etwas daran, daß jedes einfach zusammenhängende Gebiet mit mindestens zwei Randpunkten auf ein anderes solches schlichtes

[1]) Man denke sich die beiden in Aussicht genommenen Verzweigungspunkte durch eine $G_1 \cap G_2$ angehörige JORDAN-Kurve J verbunden, schneide G_1 und G_2 längs J auf und verhefte die beiden aufgeschnittenen Gebiete längs J zur zweiblättrigen RIEMANNschen Fläche.

Gebiet mit beliebig vorgegebenem Rand umkehrbar eindeutig und konform abgebildet werden kann.

Den Beweis des Satzes führe ich dadurch, daß ich zeige: Durch passende Wahl der beiden Verzweigungspunkte a und b im Durchschnitt $G_1 \cap G_2$ kann dem Modul des zweifach zusammenhängenden Gebietes jeder beliebige Wert verschafft werden. Diesen Nachweis hinwieder erbringe ich, indem ich zeige, daß der Modul eine stetige Funktion von a und b ist, daß er dem einen Ende seines Variabilitätsintervalls zustrebt, wenn a und b in einem inneren Punkt des zweifach zusammenhängenden Gebietes zusammenrücken, daß er aber dem anderen Ende seines Variabilitätsintervalls sich nähert, wenn a *oder* b einem gemeinsamen Punkt von C_1 und C_2 zustreben.

Für diese in Aussicht genommenen Nachweise ist die Definition des Moduls mit Hilfe der Methode der extremalen Länge besonders geeignet. Bildet man ein zweifach zusammenhängendes Gebiet B ohne punktförmige Randkomponente schlicht und konform auf einen Kreisring $r < |w| < 1$ ab, so ist r eine konforme Invariante der Abbildung, der sog. Modul von B. Zur Definition der extremalen Länge L betrachte man die Familie C aller in B gelegenen im Kleinen rektifizierbaren Kurven, welche die beiden Randkomponenten verbinden und lasse zur Konkurrenz des Variationsproblems

$$\frac{1}{L} = \inf_{\varrho} \iint \varrho^2 \, dx \, dy, \qquad z = x + iy$$

alle über B quadratisch integrierbaren reellwertigen Funktionen $\varrho(z)$ zu, für die $\int \varrho \, ds \geq 1$ ausfällt für jede Kurve der genannten Kurvenfamilie. L ist ebenfalls eine konforme Invariante von B. Um ihren Zusammenhang mit dem gerade genannten Modul festzustellen, bilde man B wieder auf einen Kreisring $r < |w| < 1$ schlicht konform ab. Dann liefert offenbar

$$\varrho(w) = \frac{1}{|w| \log \frac{1}{r}}$$

eine Lösung des Variationsproblems. Daher ist

$$L = \frac{1}{2\pi} \log \frac{1}{r}.$$

Nun bilde man ein gegebenes zweiblättriges RIEMANNsches Flächenstück G, das den Bedingungen des Satzes genügt, und das die beiden Verzweigungspunkte a und b hat, und das zudem beschränkt[2]) sein möge, durch

$$w = z + \sqrt{(z-a)(z-b)}$$

auf ein schlichtes zweifach zusammenhängendes Gebiet $\mathfrak{G}$ mit den beiden Randkomponenten $\mathfrak{C}_1$ und $\mathfrak{C}_2$ ab. $\mathfrak{C}_2$ möge im Inneren von $\mathfrak{C}_1$ liegen. $\mathfrak{C}_1$ und $\mathfrak{C}_2$

[2]) Es genügt diesen Fall weiterhin zu betrachten. Denn jedes andere den Bedingungen des Satzes genügende Gebiet läßt sich durch eine passende Hilfsabbildung in ein beschränktes überführen, oder kann doch mit einer gegenüber dem Text etwas modifizierten Hilfsabbildung ganz analog behandelt werden.

und damit auch $\mathfrak{G}$ hängen stetig von a und b ab. Die Definition des Moduls von $\mathfrak{G}$ vermittelst der extremalen Länge läßt aber auch erkennen, daß der Modul, d.h. also die extremale Länge stetig von a und b abhängt[3]). Diese Definition führt ferner zur Einsicht, daß die extremale Länge bei Erweiterung eines Gebietes zunimmt, mit anderen Worten, daß der innere Radius r des konzentrischen Kreisbildes $r<|w|<1$ bei Erweiterung des Gebietes abnimmt. Wenn nun a und b, die beiden Verzweigungspunkte, genügend nahe bei einem inneren Punkt a von G liegen, dann ist das schlichte durch $w=z+\sqrt{(z-a)(z-b)}$ erhaltene Gebiet $\mathfrak{G}$ mit einer inneren Randkomponente versehen, die einer Kreisscheibe um a von vorgegebener Kleinheit angehört — da nämlich der eine Zweig von $z+\sqrt{(z-a)(z-b)}$ gegen a konvergiert, wenn b gegen a rückt. Dann besitzt aber, wie eben erwähnt wurde, das konzentrische Kreisbild $r<|w|<1$ von $\mathfrak{G}$ einen inneren Kreisradius, der seinerseits beliebig klein ist. Der Modul r nähert sich also dem einen Ende 0 seines Variabilitätsintervalles $0<r<1$.

Läßt man dann bei festem a den Verzweigungspunkt b einem gemeinsamen Punkt der Ränder C_1 und C_2 der zweiblättrigen Fläche G zustreben, so erhält man durch $w=z+\sqrt{(z-a)(z-b)}$ in der w-Ebene Gebiete $\mathfrak{G}(b)$, deren beide Randkomponenten $\mathfrak{C}_1(b)$ und $\mathfrak{C}_2(b)$ (in der Nähe von $w=b$) einen gegen Null konvergierenden Abstand haben. Diese variablen Gebiete bilde man schlicht auf Kreisringe $r(b)<|w|<1$ ab und betrachte die in den Kreisringen erklärten Umkehrungsfunktionen dieser Abbildungen, die also die Ringe in die variablen Gebiete $\mathfrak{G}(b)$ mit dem Engpaß in der Nähe von $w=b$ überführen. Nun nehme man an, daß nicht $r(b)\to 1$ gilt, wenn b sich einem gemeinsamen Punkt von C_1 und C_2 nähert. Dann gibt es einen allen Ringen gemeinsamen Durchschnittsring, sagen wir $R<|w|<1$, in dem die genannten Umkehrungsfunktionen eindeutig schlicht und gleichmäßig beschränkt sind. Man wähle eine in $R_1<|w|<R_2$, $R_1=R+\frac{1}{4}(1-R)$, $R_2=1-\frac{1}{4}(1-R)$ gleichmäßig konvergente Teilfolge aus. Ihre Grenzfunktion kann nicht konstant sein. Denn alle Glieder der gleichmäßig konvergenten Folge bilden den Kreis $|w|=\frac{1+R}{2}$ auf Kurven ab, die die innere Randkomponente $\mathfrak{C}_2(b)$ von $\mathfrak{G}(b)$ umschließen. Deren Durchmesser konvergiert aber nicht gegen 0, wenn b sich einem Schnittpunkt von C_1 und C_2 nähert, während a festgehalten wird. Wäre die Grenzfunktion konstant, so müßten bekanntlich die Ableitungen der Glieder der Folge gegen 0 streben, was der eben genannten Erscheinung einer Bildkurve widerspricht, deren Länge nicht nach 0 strebt. Ist aber die Grenzfunktion nicht konstant, so bildet sie bekanntlich den Ring $R_1<|w|<R_2$ schlicht ab. Der absolute Betrag der Ableitung der Grenzfunktion in den Punkten des Kreises $|w|=\frac{1+R}{2}$ liegt daher oberhalb einer positiven Schranke. Daher ist das auch für die Glieder der konvergenten Folge bei genügend großer

[3]) Der Beweis gelingt auch mit Hilfe einer passenden Verallgemeinerung des Carathéodoryschen Begriffs „Kern einer Gebietsmenge" auf zweifach zusammenhängende Gebiete.

Gliednummer ebenso. Daher geht nach dem KOEBEschen Verzerrungssatz der Kreisring $R_1 < |w| < R_2$ in $\mathfrak{C}_2(b)$ umschließende zweifach zusammenhängende Bereiche über, deren Ränder für alle Glieder der Folge voneinander einen Abstand oberhalb einer positiven Schranke haben. Das widerspricht der Existenz des erwähnten Engpasses.

Damit ist der Beweis des Satzes abgeschlossen. Es sei noch folgende Bemerkung angefügt. Die Voraussetzung, daß der Durchschnitt $C_1 \cap C_2$ nicht leer ist, ist wesentlich. Denn läge C_2 ganz im Inneren von C_1, so wäre der Modul r eines jeden zweiblättrigen aus G_1 und G_2 aufgebauten Gebietes G kleiner als der Modul r des zweifach zusammenhängenden schlichten Gebietes, das zwischen C_1 und C_2 eingespannt ist, so daß also derartige Gebiete nicht das ganze Modulintervall ausschöpfen können.

Offenbar genügt es laut Beweisgang für die Gültigkeit des Satzes, wenn man a festhält und b auf einer festgewählten stetigen Kurve aus $G_1 \cap G_2$ annimmt, die a mit einem gemeinsamen Punkt von C_1 und C_2 verbindet, wobei natürlich diese Kurven in ihrer Projektion auf die z-Ebene gemeint sind.

Die Frage der Unität bedarf noch weiterer Überlegung, ebenso wie der Wunsch, analoge Sätze auch für mehr als zweifach zusammenhängende Gebiete auszusprechen und zu begründen. Es kam mir aber in dieser Note nur darauf an, die Aufmerksamkeit auf einen Sachverhalt zu lenken, der mir recht merkwürdig erscheint.

Literatur

[1] FABER, G.: Über den Hauptsatz der Theorie der konformen Abbildung. Sitzgsber. Bayr. Akad. Wiss., Math.-phys. Kl. **1922**, 91—100. — [2] BIEBERBACH, L.: Über einen Riemannschen Satz aus der Lehre von der konformen Abbildung. Sitzgsber. Berl. math. Ges. **1925**, 6—9. — [3] GRUNSKY, H.: Über die konforme Abbildung mehrfach zusammenhängender Bereiche auf mehrblättrige Kreise. Sitzgsber. Pr. Akad. Wiss., Phys.-math. Kl. **1937**, 40—46; II **1941**, Nr. 11, 1—8.

Berlin-Dahlem, Gelfertstr. 16

(Eingegangen am 21. April 1956)

GRAUERT, H., u. R. REMMERT
Math. Zeitschr. Bd. 67, S. 103—128 (1957)

Singularitäten komplexer Mannigfaltigkeiten und Riemannsche Gebiete

Herrn GEORG FABER zum 80. Geburtstag am 5. April 1957 gewidmet

Von

HANS GRAUERT und REINHOLD REMMERT

Einleitung

1. In der algebraischen Geometrie hat man seit jeher den Singularitäten der algebraischen Mannigfaltigkeiten ein besonderes Interesse entgegengebracht (s. z.B. [7], [27], [41]). Es zeigte sich bald, daß die Singularitäten in zwei Klassen zerfallen. Einmal gibt es Punkte, in denen die Singularität lediglich durch die Art der Einbettung der algebraischen Fläche in den projektiven Raum bedingt ist; in diesem Falle kann man stets eine Umgebung des singulären Punktes holomorph auf eine Hyperkugel abbilden. Zum anderen fand man aber auch Punkte, die nicht in dem soeben beschriebenen Sinne uniformisierbar sind. Es ist ein bekanntes Problem der algebraischen Geometrie, diese nicht uniformisierbaren Singularitäten durch birationale Transformationen zu eliminieren[1]).

In der Funktionentheorie mehrerer Veränderlichen hat man im Gegensatz zur algebraischen Geometrie lange Zeit gezögert, singuläre Punkte in den Kreis der Betrachtungen einzubeziehen. Die gesamte Theorie der Holomorphiegebiete ist z.B. im Ergebnisbericht von BEHNKE und THULLEN [2] lediglich für die schlichten und unverzweigten Gebiete über dem Raum C^n von n komplexen Veränderlichen entwickelt worden; verzweigte Gebiete wurden nicht betrachtet. Bis heute fehlt die Übertragung auf beliebige Gebiete über dem C^n.

Erst in der neueren Zeit ist man auch in der komplexen Analysis dazu übergegangen, verzweigte Gebiete — sog. RIEMANNsche Gebiete — zuzulassen und die algebroiden Funktionselemente zu solchen Gebilden zusammenzufassen. Dabei zeigte sich, daß auch hier nicht-uniformisierbare Punkte auftreten. In einer Arbeit von BEHNKE und STEIN [5] wird explizit das analytische Gebilde der Funktion $\sqrt{z_1 z_2}$ als verzweigtes Gebiet über dem C^2 aufgewiesen, das über dem Nullpunkt einen nicht-uniformisierbaren Punkt besitzt.

2. Wie in der klassischen Funktionentheorie hat man auch in der Funktionentheorie mehrerer Veränderlichen neben den konkreten Gebieten über

[1]) Für zwei- sowie dreidimensionale algebraische Mannigfaltigkeiten ist bekannt, daß sich die Singularitäten durch birationale Transformationen auflösen lassen. Vergleiche hierzu etwa [24], [25], [43]. Für den Fall höherer Dimension konnte das Problem bis heute noch nicht gelöst werden.

dem C^n abstrakte Gebilde definiert. Nach dem Vorbild der allgemeinen Topologie der Mannigfaltigkeiten wurden aber auch hier zunächst die singulären Punkte ausgeschlossen. Die von CARATHÉODORY [8], TEICHMÜLLER [38] und HOPF [26] definierten komplexen Mannigfaltigkeiten enthalten nur uniformisierbare Punkte[2]). Inzwischen ist man jedoch auch in der Funktionentheorie auf abstrakten Gebilden dazu übergegangen, die nicht-uniformisierbaren Singularitäten als gleichberechtigt anzusehen. Das führte BEHNKE und STEIN [5] sowie CARTAN [13], [15] zur Definition des komplexen Raumes (espace analytique général). Darunter versteht man einen HAUSDORFFschen Raum, dem in bestimmter Weise eine komplex-analytische Struktur aufgeprägt ist. Komplexe Räume haben folgende Eigenschaft: Jeder Punkt besitzt eine Umgebung, die zu einem sog. analytischen Repräsentanten analytisch äquivalent ist. Dabei ist ein analytischer Repräsentant bei BEHNKE und STEIN eine endlich-blättrige, analytisch-verzweigte Überlagerung einer offenen Hyperkugel H im C^n; analytisch-verzweigt heißt, daß die Grundpunktmenge der Verzweigungspunkte das genaue Nullstellengebilde einer in H nicht identisch verschwindenden holomorphen Funktion ist (zur genauen Beschreibung vgl. § 1, 1—4). CARTAN dagegen verwendet als analytische Repräsentanten sog. Überlagerungsräume von analytischen Mengen (s. § 2, 1—3). Es ist bislang noch ein offenes Problem, ob beide Definitionen äquivalent sind.

3. Über die Funktionentheorie in verzweigten Gebieten über dem C^n sowie auch in komplexen Räumen ist bislang nur wenig bekannt geworden. In der vorliegenden Arbeit wird an Hand eines Beispiels gezeigt, daß in der Umgebung einer nicht-uniformisierbaren Stelle eines komplexen Raumes im allgemeinen Aussagen gelten, die in keiner Analogie zu den Sätzen stehen, die von der komplexen Analysis auf komplexen Mannigfaltigkeiten her bekannt sind. So zeigen wir in § 2:

I. *Eine rein $(n-1)$-dimensionale analytische Menge in einem n-dimensionalen komplexen Raum ist lokal nicht stets das genaue Nullstellengebilde einer lokal-holomorphen Funktion* $(n \geq 3)$.

II. *Zwei rein $(n-1)$-dimensionale analytische Mengen in einem n-dimensionalen komplexen Raum können sich punkthaft schneiden* $(n \geq 3)$. *Die Schnittmenge braucht also nicht wie in komplexen Mannigfaltigkeiten eine rein $(n-2)$-dimensionale analytische Menge zu sein.*

III. *Es gibt komplexe Räume beliebig großer Dimension mit meromorphen Funktionen, deren Unbestimmtheitsstellen isoliert liegen.*

In § 3 und 4 untersuchen wir verzweigte RIEMANNsche Gebiete über dem C^n. OKA hat 1953 bewiesen [30], daß jedes unverzweigte Holomorphiegebiet holomorph-konvex ist. Ferner zeigte er, daß die unverzweigten Holomorphiegebiete durch eine lokale Randeigenschaft charakterisiert werden können. Er bewies nämlich, daß jedes unverzweigte RIEMANNsche Gebiet über dem C^n

[2]) Der Ausdruck „*komplexe Mannigfaltigkeit*" wird unseres Wissens erstmals in [26] benutzt; in [8] sowie [38] handelt es sich indessen um die gleiche Begriffsbildung.

genau dann ein Holomorphiegebiet ist, wenn jeder Randpunkt ein sog. pseudokonvexer Randpunkt ist. Wir zeigen, daß diese Sätze für verzweigte RIEMANNsche Gebiete über dem C^n im allgemeinen falsch sind, wenn $n \geq 3$. Es ergibt sich nämlich:

IV. *Es gibt endlich-blättrige, n-dimensionale Holomorphiegebiete, die aus lauter uniformisierbaren Punkten bestehen und mindestens einen nicht pseudokonvexen Randpunkt besitzen,* $n \geq 3$.

V. *Es gibt endlich-blättrige, n-dimensionale Holomorphiegebiete, die aus lauter uniformisierbaren Punkten bestehen und nicht holomorph-konvex sind,* $n \geq 3$.

Aus V. folgt insbesondere, daß die verzweigten Holomorphiegebiete nicht gegenüber umkehrbar holomorphen Abbildungen invariant sind. Für unverzweigte RIEMANNsche Gebiete ist bekanntlich die Eigenschaft, Holomorphiegebiet zu sein, eine analytische Invariante (vgl. z.B. [12], p. 154).

§ 1. Komplexe Räume

1. Betrachtet man eine n-dimensionale algebraische Mannigfaltigkeit in der Umgebung einer Singularität [z.B. $w^q - z_1 z_2 \dots z_n = 0, q > 1, n > 1$, im Nullpunkt[3])], so stellt man fest, daß diese dort stets die Struktur einer endlich-blättrigen, analytisch-verzweigten Überlagerung besitzt. Es werde daher zunächst dieser Begriff definiert (vgl. auch [21], [23]).

Def. 1 *(Analytisch-verzweigte Überlagerung). Ein Tripel* $\mathfrak{R} = (R, \Phi, G)$ *heißt eine (n-dimensionale) endlich-blättrige, analytisch-verzweigte Überlagerung eines Gebietes G des n-dimensionalen komplexen Zahlenraumes* C^r, *wenn folgendes gilt:*

0) *R ist ein* HAUSDORFFscher *Raum;* Φ *ist eine stetige, eigentliche, nirgends entartete Abbildung*[4]) *von R in G.*

1) *Es gibt eine analytische Menge*[5]) A *in G, derart, daß* $\Phi^{-1}(A)$ *den Raum R nirgends zerlegt und die Menge* $R - \Phi^{-1}(A)$ *durch* Φ *lokal-topologisch in G abgebildet wird.*

Dabei sagen wir, daß eine Teilmenge S eines topologischen Raumes R den Raum R *nirgends zerlegt*, wenn S nirgends dicht[6]) in R liegt und für jede

[3]) Vgl. etwa [5], p. 4, wo diese algebraische Mannigfaltigkeit für den Fall $n = q = 2$ ausführlich diskutiert wird.

[4]) Eine stetige Abbildung φ eines topologischen Raumes R in einen topologischen Raum R' heißt *eigentlich*, wenn das φ-Urbild jeder kompakten Teilmenge von R' eine kompakte Menge in R ist. φ heißt *nirgends entartet*, wenn das φ-Urbild eines jeden Punktes von R' eine diskrete Menge in R ist.

[5]) Eine Teilmenge M eines Gebietes G des Zahlenraumes C^n heißt eine *analytische Menge* in G, wenn es zu jedem Punkt $\mathfrak{z}_0 \in G$ eine Umgebung $U(\mathfrak{z}_0)$ und endlich viele in $U(\mathfrak{z}_0)$ holomorphe Funktionen $f_1(\mathfrak{z}), \dots, f_s(\mathfrak{z})$ gibt, derart, daß ein Punkt $\mathfrak{z}' \in U(\mathfrak{z}_0)$ genau dann zu M gehört, wenn gilt: $f_1(\mathfrak{z}') = \dots = f_s(\mathfrak{z}') = 0$. Zur Theorie der analytischen Mengen vgl. etwa [13], [15], [23], [33], [35].

[6]) Eine Teilmenge S eines topologischen Raumes R liegt *nirgends dicht* in R, wenn jede nichtleere offene Menge U von R eine nichtleere offene Menge V enthält, die keinen Punkt mit S gemeinsam hat.

offene zusammenhängende Teilmenge $U \subset R$ gilt, daß $U - U \cap S$ zusammenhängend[7]) ist.

Aus Def. 1 folgt sofort: *Ist $\mathfrak{R} = (R, \Phi, G)$ eine analytisch-verzweigte Überlagerung, so ist der Raum R lokal-zusammenhängend. Die Menge $R - \Phi^{-1}(A)$ liegt dicht in R.*

Ein Punkt $r \in R$ heißt über dem Punkt $\mathfrak{z} = \Phi(r) \in G$ gelegen. Da Φ eigentlich und nirgends entartet abbildet, so liegen über jedem Punkt $\mathfrak{z} \in G$ nur endlich viele Punkte von R. Man kann zeigen, daß über jedem Punkt $\mathfrak{z} \in G - A$ gleich viel Punkte von R liegen; ihre Anzahl heißt die *Blätterzahl* der Überlagerung.

Jedes Gebiet $G \subset C^n$ besitzt die sog. triviale einblättrige Überlagerung (G, ι, G), wo ι: $G \to G$ die Injektion bezeichnet.

Die Punkte von R, in denen Φ nicht lokal-topologisch ist, heißen *Verzweigungspunkte* von R.

Es ist evident, daß R außerhalb der Verzweigungspunkte sich *lokal sphärisch* verhält: in der Umgebung jedes dieser Punkte ist R der $2n$-dimensionalen Zelle homöomorph. In den Verzweigungspunkten ist das im allgemeinen, falls $n \geq 2$, nicht mehr der Fall [vgl. z. B. die durch $w^2 - z_1 z_2 = 0$ definierte Überlagerung im Punkte über $(0, 0)$]. Man kann jedoch noch folgendes zeigen (vgl. [23]):

Über den gewöhnlichen Punkten[8]) der analytischen Menge $A \subset G$ liegen nur Windungspunkte von $\mathfrak{R}$.

Dabei heißt ein Punkt $r \in R$ ein *Windungspunkt* von $\mathfrak{R}$, wenn es eine Umgebung $U(r)$ und eine topologische Abbildung ψ von U auf ein Gebiet W des Raumes der Veränderlichen $w_1, \ldots, w_n$ gibt, derart, daß $\Phi \circ \psi^{-1}$ sich bei geeigneter Numerierung der Koordinaten $z_1, \ldots, z_n$ in der Form

$$z_1 = w_1, \ldots, z_{n-1} = w_{n-1}, z_n = w_n^q + g(w_1, \ldots, w_{n-1})$$

mit einer holomorphen Funktion $g(w_1, \ldots, w_{n-1})$ schreiben läßt.

Alle Windungspunkte einer analytisch-verzweigten Überlagerung $\mathfrak{R}$ sind offensichtlich sphärische Punkte von $\mathfrak{R}$.

2. Die analytisch-verzweigten Überlagerungen sollen zur lokalen Charakterisierung der komplexen Räume herangezogen werden. Dazu benötigen wir jedoch noch den grundlegenden Begriff der holomorphen Abbildung einer

---

[7]) Wir wollen eine Teilmenge S eines topologischen Raumes R *zusammenhängend* nennen, wenn zwei beliebige Punkte $r_1, r_2 \in S$ stets durch einen in S verlaufenden Weg (Kurve) verbindbar sind. Unter einem Weg in R verstehen wir eine stetige Abbildung $\{r = r(t), 0 \leq t \leq 1\}$ des Einheitsintervalles $0 \leq t \leq 1$ in den Raum R.

[8]) Ein Punkt $\mathfrak{z}_0$ einer analytischen Menge A heißt ein *gewöhnlicher Punkt* von A, wenn A in einer Umgebung von $\mathfrak{z}_0$ durch k holomorphe Funktionen $f_1, \ldots, f_k$ beschrieben werden kann, deren Funktionalmatrix $(\partial f_\varkappa / \partial z_\nu)$, $\varkappa = 1, \ldots, k$; $\nu = 1, \ldots, n$, in $\mathfrak{z}_0$ vom Range k ist.

analytisch-verzweigten Überlagerung $\mathfrak{R}$ in eine ebensolche Überlagerung $\mathfrak{R}'$. Vorbereitend definieren wir:

Def. 2 *(Holomorphe Funktion in einer analytisch-verzweigten Überlagerung). Eine stetige komplex-wertige Funktion f in einer offenen Menge $W \subset R$ heißt eine holomorphe Funktion in W, wenn es zu jedem Punkt $r \in W - W \cap \Phi^{-1}(A)$ eine Umgebung $U(r) \subset W$ gibt, die durch Φ topologisch auf $\Phi(U(r))$ abgebildet wird, derart, daß $f \circ \Phi^{-1}$ eine holomorphe Funktion in $\Phi(U(r))$ ist.*

Nunmehr läßt sich der Begriff der holomorphen Abbildung einführen:

Def. 3 *(Holomorphe Abbildung einer analytisch-verzweigten Überlagerung) Sind $\mathfrak{R} = (R, \Phi, G)$ und $\mathfrak{R}' = (R', \Phi', G')$ zwei analytisch-verzweigte Überlagerungen, so heißt eine stetige Abbildung τ einer offenen Menge $W \subset R$ in R' eine holomorphe Abbildung, wenn folgendes gilt: Ist $f(r')$ eine holomorphe Funktion in einer offenen Menge $W' \subset R'$, so ist $f(\tau(r))$ eine holomorphe Funktion in der offenen Menge $\tau^{-1}(W') \subset W$.*

Auf Grund dieser Definition sind die holomorphen Funktionen auf $\mathfrak{R}$ genau die holomorphen Abbildungen von $\mathfrak{R}$ in die Zahlenebene C^1.

3. Komplexe Räume sind topologische Räume, die mit einer sog. komplex-analytischen Struktur versehen sind. Der Begriff der komplex-analytischen Struktur läßt sich besonders elegant fassen, wenn man sich einer von EHRESMANN [17], [18] eingeführten Terminologie bedient (vgl. auch [32]).

Def. 4 *(Komplexe Karte). Eine n-dimensionale, komplexe Karte auf einem* HAUSDORFF*schen Raum X ist ein Paar (U, ψ), wobei U eine offene zusammenhängende Menge von X und ψ eine topologische Abbildung von U auf eine offene Menge einer n-dimensionalen analytisch-verzweigten Überlagerung ist.*

Zwischen komplexen Karten wird eine „Verträglichkeitsbeziehung" definiert:

Def. 5 *(Holomorph verträgliche komplexe Karten). Zwei komplexe Karten (U_1, ψ_1) und (U_2, ψ_2) auf einem* HAUSDORFF*schen Raum X heißen holomorph verträglich, wenn $U_1 \cap U_2$ leer ist oder wenn die Abbildung $\psi_2 \circ \psi_1^{-1}$ eine umkehrbar holomorphe Abbildung von $\psi_1(U_1 \cap U_2)$ auf $\psi_2(U_1 \cap U_2)$ ist.*

Komplexe Karten werden nun zu komplexen Atlanten zusammengefaßt.

Def. 6 *(Komplexer Atlas). Ein komplexer Atlas $\mathfrak{A}$ auf einem* HAUSDORFF*schen Raum X ist eine Kollektion von paarweise miteinander holomorph verträglichen komplexen Karten $\{(U_\iota, \psi_\iota), \iota \in J\}$, wobei $\bigcup_{\iota \in J} U_\iota = X$.*

Ein komplexer Atlas $\mathfrak{A}$ auf X heißt *vollständig*, wenn jede komplexe Karte (V, ψ) auf X, die mit allen Karten aus $\mathfrak{A}$ holomorph verträglich ist, zu $\mathfrak{A}$ gehört.

Man zeigt leicht, daß sich jeder komplexe Atlas vervollständigen läßt.

Besitzt ein HAUSDORFFscher Raum X einen komplexen Atlas $\mathfrak{A}$, so ist X im Kleinen einer analytisch-verzweigten Überlagerung homöomorph und somit lokal-zusammenhängend. Daher zerfällt X in zusammenhängende Komponenten $X_\varkappa$, $\varkappa \in \mathsf{K}$. Eine Karte $(U, \psi) \in \mathfrak{A}$ ist stets, da U zusammenhängend ist, eine komplexe Karte auf genau einer solchen Komponente.

Da man zwei Karten aus $\mathfrak{A}$ auf einer Komponente $X_\varkappa$ stets durch eine endliche Kette von Karten aus $\mathfrak{A}$ miteinander verbinden kann und zwei benachbarte Karten (U_1, ψ_1) und (U_2, ψ_2) stets dieselbe Dimension[9]) besitzen (die Abbildung $\psi_2 \circ \psi_1^{-1}$ ist topologisch!), haben alle Karten, die zur selben Komponente $X_\varkappa$ gehören, die gleiche Dimension. Daher kommt jeder Komponente $X_\varkappa$ eine eindeutig bestimmte, *endliche Dimension* $d(X_\varkappa)$ zu.

Die Dimension von X selbst wird nun definiert als: $d(X) = \sup_{\varkappa \in K} d(X_\varkappa)$. $d(X)$ kann unendlich sein.

Haben alle Komponenten $X_\varkappa$ von X die gleiche Dimension d, so nennen wir X *rein d-dimensional*; wir schreiben dann auch: $X = X^d$.

4. Wir sind nun in der Lage, die Begriffe der komplex-analytischen Struktur, des komplexen Raumes usw. einzuführen.

Def. 7 *(Komplex-analytische Struktur). Eine komplex-analytische Struktur in einem* HAUSDORFF*schen Raum X ist ein vollständiger komplexer Atlas auf X.*

Def. 8 *(Komplexer Raum). Ein* HAUSDORFF*scher Raum X mit einer komplex-analytischen Struktur heißt ein komplexer Raum.*

Offenbar kann der Träger R jeder analytisch-verzweigten Überlagerung $\mathfrak{R} = (R, \Phi, G)$ als komplexer Raum aufgefaßt werden, indem man den auf R einelementrigen komplexen Atlas (R, ι), wo ι die Injektion ist, vervollständigt.

Def. 9 *(Uniformisierbarer Punkt). Ein Punkt x eines komplexen Raumes X heißt ein uniformisierbarer Punkt von X, wenn es eine Karte $(U, \psi) \in \mathfrak{A}$ gibt, derart, daß $x \in U$ und ψ die Menge U auf die triviale Überlagerung eines Gebietes des C^n abbildet.*

Die komplexen Mannigfaltigkeiten[10]) sind nun spezielle komplexe Räume.

Def. 10 *(Komplexe Mannigfaltigkeit). Eine komplexe Mannigfaltigkeit ist ein komplexer Raum, der nur aus uniformisierbaren Punkten besteht.*

In jedem Teilbereich (= offene Menge) $\tilde{X}$ eines komplexen Raumes (bzw. einer komplexen Mannigfaltigkeit) X wird in natürlicher Weise durch die komplex-analytische Struktur von X eine komplex-analytische Struktur induziert, die $\tilde{X}$ zu einem komplexen Raum (bzw. einer komplexen Mannigfaltigkeit) macht. Ferner ist das topologische Produkt zweier komplexer Räume (bzw. zweier komplexer Mannigfaltigkeiten) in natürlicher Weise mit einer komplex-analytischen Struktur versehen und mithin ein komplexer Raum (bzw. eine komplexe Mannigfaltigkeit).

5. In komplexen Räumen läßt sich der Begriff der holomorphen Funktion einführen.

Def. 11 *(Holomorphe Funktion in einem komplexen Raum). Eine komplexwertige Funktion $f(x)$ in einem komplexen Raum X heißt holomorph in X, wenn für jede komplexe Karte $(U, \psi) \in \mathfrak{A}$ die Funktion $f \circ \psi^{-1}$ eine in $\psi(U)$ holomorphe Funktion ist.*

[9]) Nach K. MENGER [28] kommt jedem topologischen Raum R eine wohldefinierte Dimension zu. Dieselbe ist invariant gegenüber topologischen Abbildungen. Vgl. auch HUREWICZ, W., and H. WALLMANN: Dimension Theory, p. 24 ff. Princeton 1948.

[10]) Zur Def. der komplexen Mannigfaltigkeiten vgl. auch [26].

Man überlegt sich sofort:

Eine in einem komplexen Raum X gegebene komplex-wertige Funktion $f(x)$ ist bereits dann holomorph in X, wenn es zu jedem Punkt $x \in X$ eine komplexe Karte $(U, \psi) \in \mathfrak{A}$ mit $x \in U$ gibt, so daß $f \circ \psi^{-1}$ in $\psi(U)$ holomorph ist.

Hieraus folgt:

Ist $\mathfrak{R} = (R, \Phi, G)$ eine analytisch-verzweigte Überlagerung, so ist eine auf R komplex-wertige Funktion f genau dann holomorph auf R im Sinne von Def. 2, wenn sie auf R im Sinne von Def. 11 holomorph ist (wobei R als komplexer Raum aufzufassen ist).

Man überzeugt sich direkt, daß die Summe sowie das Produkt zweier holomorpher Funktionen wieder eine holomorphe Funktion ist. Wegen des Identitätssatzes für holomorphe Funktionen treten keine Nullteiler bezüglich der Multiplikation auf, wenn der Raum zusammenhängend ist. Daher folgt:

Die Gesamtheit aller in einem komplexen Raum X holomorphen Funktionen bildet einen Ring $R(X)$. Der Ring $R(X)$ ist genau dann ein Integritätsring, wenn der Raum X zusammenhängend ist.

Der Begriff der holomorphen Funktion läßt sich verallgemeinern zum Begriff der holomorphen Abbildung.

Def. 12 *(Holomorphe Abbildung eines komplexen Raumes). Eine stetige Abbildung $\tau: X \to Y$ eines komplexen Raumes X in einen komplexen Raum Y heißt eine holomorphe Abbildung, wenn folgendes gilt: Ist $f(y)$ irgendeine holomorphe Funktion in einem Teilbereich Y^* von Y, so ist $f(\tau(x))$ eine holomorphe Funktion im Bereich $\tau^{-1}(Y^*) \subset X$.*

Aus dieser Definition folgt unmittelbar:

Eine eindeutige Abbildung τ eines komplexen Raumes X in den Zahlenraum C^n ist genau dann holomorph, wenn die τ beschreibenden n Abbildungsfunktionen $z_1 = f_1(x), \ldots, z_n = f_n(x)$, $x \in X$, holomorphe Funktionen in X sind.

Weiter ergibt sich:

Eine Abbildung τ einer analytisch-verzweigten Überlagerung $\mathfrak{R} = (R, \Phi, G)$ in eine analytisch-verzweigte Überlagerung $\mathfrak{R}' = (R', \Phi', G')$ ist genau dann holomorph im Sinne von Def. 3, wenn sie holomorph im Sinne von Def. 12 ist (wobei R und R' als komplexe Räume aufzufassen sind).

6. Den Nullstellenmengen holomorpher Funktionen kommt ein besonderes Interesse zu. Wir definieren daher:

Def. 13 *(Analytische Menge). Eine Teilmenge M eines komplexen Raumes X heißt eine analytische Menge in X, wenn es zu jedem Punkt $x' \in X$ eine Umgebung U und endlich viele in U holomorphe Funktionen $f_1, \ldots, f_s$ gibt, so daß ein Punkt $x \in U$ genau dann zu M gehört, wenn gilt: $f_1(x) = \cdots = f_s(x) = 0$.*

Offenbar ist eine analytische Menge immer eine abgeschlossene Menge.

Jede analytische Menge M in einem komplexen Raum X hat in jedem ihrer Punkte x eine wohlbestimmte *Dimension* d'_x [11]). Ein Satz sagt aus, daß stets gilt: $d'_x \equiv 0 (2)$. Wir setzen $d_x = \frac{1}{2} d'_x$ und nennen d_x die *(komplexe)*

[11]) Vgl. Fußnote [9]).

Dimension von M in x. Faßt man X als analytische Menge in X auf, so stimmt diese Dimension von X in x mit der in 3. definierten Dimension derjenigen zusammenhängenden Komponente von X überein, die x enthält.

Wir benötigen im folgenden:

Ist M eine analytische Menge in einem komplexen Raum X, die keine zusammenhängende Komponente von X enthält, so zerlegt M den Raum X nirgends[12]).

Der Begriff der analytischen Menge wird verallgemeinert zum Begriff der (analytisch) dünnen Menge.

Def. 14 *(Dünne Menge). Eine abgeschlossene Teilmenge D eines komplexen Raumes X heißt eine dünne Menge von der Ordnung k $(1 \leq k \leq d(X))$, wenn es zu jedem Punkt $x \in D$ eine Umgebung U und eine in U analytische Menge M gibt, so daß gilt: $D \cap U \subset M$, $d_x(M) \leq d_x(X) - k$. — Die leere Menge sei dünn von der Ordnung $d(X) + 1$.*

Ein leicht zu beweisender Satz ist nun (vgl. [*23*]):

Die nicht-uniformisierbaren Punkte eines komplexen Raumes bilden eine dünne Menge der Ordnung 2[13]).

7. Neben holomorphen Funktionen betrachten wir auch sog. *meromorphe Funktionen*. Wir sagen, in einem komplexen Raum X ist eine meromorphe Funktion $h(x)$ gegeben, wenn folgende Situation vorliegt:

α) Außerhalb einer in X dünnen Menge P erster Ordnung ist eine holomorphe Funktion $h^*(x)$ gegeben, die in keinen Punkt von P holomorph fortsetzbar ist.

β) Zu jedem Punkt $x \in P$ gibt es eine Umgebung U und eine in U holomorphe, nicht identisch verschwindende Funktion $f(x)$, so daß die in $U - P$ holomorphe Funktion $f(x) \cdot h^*(x)$ in ganz U holomorph fortsetzbar ist.

Die Menge P heißt die *Polstellenmenge* von $h(x)$; man kann zeigen, daß P sogar eine analytische Menge ist[14]).

Eine in einem komplexen Raum X meromorphe Funktion $h(x)$ gibt zu einer holomorphen Abbildung von $X - P$ in die Zahlenebene Anlaß. Faßt man diese Abbildung als Abbildung in die Zahlenkugel auf, so läßt sie sich stets in eine nichtleere Teilmenge P^* von P eindeutig und holomorph fortsetzen. Die Punkte von P^* heißen *außerwesentliche Singularitäten 1. Art* oder *eigentliche Polstellen* der meromorphen Funktion $h(x)$; die Punkte von $P - P^*$ heißen *außerwesentliche Singularitäten 2. Art* oder *Unbestimmtheitsstellen* der meromorphen Funktion $h(x)$. Man kann zeigen, daß $P - P^*$ eine dünne (sogar analytische) Menge 2. Ordnung in X ist[15]).

Für meromorphe Funktionen läßt sich in natürlicher Weise eine Addition, Multiplikation und Division erklären. Es ergibt sich:

Die Gesamtheit $K(X)$ aller in einem zusammenhängenden komplexen Raum X meromorphen Funktionen bildet einen Körper.

---

[12]) Zum Beweise vgl. [*23*].

[13]) Man kann sogar zeigen, daß die nichtuniformisierbaren Punkte eine analytische Menge bilden, vgl. [*15*], Exp. X u. XI.

[14]) Vgl. etwa [*2*] sowie [*23*].

[15]) Vgl. hierzu [*34*].

§ 2. Komplexe Überlagerungsräume analytischer Mengen. Untersuchung eines speziellen komplexen Raumes mit einer Singularität

1. Ist X eine komplexe Mannigfaltigkeit und M eine analytische Menge in X, die nur aus gewöhnlichen Punkten besteht, so ist M in natürlicher Weise mit einer komplex-analytischen Struktur versehen, die M selbst zu einer komplexen Mannigfaltigkeit macht; die Injektion $\iota: M \to X$ ist dabei eine holomorphe Abbildung. Dagegen tragen analytische Mengen mit nichtgewöhnlichen Punkten im allgemeinen keine natürliche komplex-analytische Struktur. Es gibt im C^n $(n \geq 3)$ sogar analytische Mengen M, die in einem Punkt $\mathfrak{z}_0$ irreduzibel sind und denen in keiner Weise in einer Umgebung $U(\mathfrak{z}_0) < M$ eine komplex-analytische Struktur $\mathfrak{A}$ aufgeprägt werden kann, derart, daß die Injektion $\iota: U(\mathfrak{z}_0) \to C^n$ eine holomorphe Abbildung ist (vgl. § 2.7)[16]. Jeder analytischen Menge in einer komplexen Mannigfaltigkeit läßt sich jedoch in eindeutiger Weise ein komplexer Raum zuordnen. Wir definieren:

Def. 15 *(Komplexer Überlagerungsraum). Ist X eine komplexe Mannigfaltigkeit und M eine analytische Menge in X, so heißt ein Paar (M^*, μ) ein komplexer Überlagerungsraum von M, wenn die folgenden Bedingungen erfüllt sind:*

0) *M^* ist ein komplexer Raum; μ ist eine eigentliche, nirgends entartete holomorphe Abbildung von M^* in X mit $\mu(M^*) = M$.*

1) *Bezeichnet N die Menge der nichtgewöhnlichen Punkte von M, so enthält $\mu^{-1}(N)$ keine ganze zusammenhängende Komponente des Raumes M^*; die Menge $M^* - \mu^{-1}(N) = \overset{\circ}{M}{}^*$ wird durch μ umkehrbar holomorph auf die komplexe Mannigfaltigkeit $M - N = \overset{\circ}{M}$ abgebildet*[17]).

Die nichtgewöhnlichen Punkte N einer in einer komplexen Mannigfaltigkeit X analytischen Menge M bilden selbst wieder eine analytische Menge in X, deren Dimension in jedem Punkte $x \in N$ kleiner als die Dimension von M in x ist (vgl. [15], Exp. IX). Da die Urbilder analytischer Mengen bezüglich holomorpher Abbildungen wieder analytische Mengen sind, so folgt wegen der in 1.6 gemachten Bemerkung aus § 1:

Der komplexe Raum M^* wird durch die Menge $\mu^{-1}(N)$ nirgends zerlegt.

2. Wir beweisen in diesem Abschnitt die Einzigkeit komplexer Überlagerungsräume.

Satz 1. *Eine in einer komplexen Mannigfaltigkeit X analytische Menge M besitzt höchstens einen komplexen Überlagerungsraum.* [*Das heißt: sind (M_1^*, μ_1)*

[16]) Natürlich ist hier der Begriff der komplex-analytischen Struktur im Sinne von Def. 7 zu verstehen. In [33] ist der Begriff des komplexen Raumes so gefaßt, daß eine analytische Menge in X stets in natürlicher Weise als komplexer Unterraum von X angesehen werden kann.

[17]) Bereits K. STEIN [36] hat eine begriffliche Charakterisierung des komplexen Überlagerungsraumes gegeben. Die hier den Überlagerungen zugrunde gelegte Definition stimmt nicht mit der STEINschen Definition überein.

und (M_2^*, μ_2) zwei komplexe Überlagerungsräume von M, so gibt es eine umkehrbar holomorphe Abbildung τ von M_1^* auf M_2^*, so daß gilt: $\mu_1 = \mu_2 \circ \tau$.]

Beweis. Seien (M_1^*, μ_1) und (M_2^*, μ_2) zwei komplexe Überlagerungsräume von M. Da μ_1 bzw. μ_2 den Raum $\overset{\circ}{M}{}_1^*$ bzw. $\overset{\circ}{M}{}_2^*$ umkehrbar holomorph auf $\overset{\circ}{M}$ abbildet, wird durch $\overset{\circ}{\tau}(m_1) = \mu_2^{-1} \circ \mu_1(m_1)$, $m_1 \in \overset{\circ}{M}{}_1^*$, eine umkehrbar holomorphe Abbildung von $\overset{\circ}{M}{}_1^*$ auf $\overset{\circ}{M}{}_2^*$ definiert. Diese Abbildung $\overset{\circ}{\tau}$ werde nun zu einer stetigen Abbildung τ von M_1^* in M_2^* fortgesetzt! Ist $m_1 \in M_1^* - \overset{\circ}{M}{}_1^*$ irgendein Punkt, so sei $m_1^{(\nu)}$ eine gegen m_1 konvergierende Folge aus $\overset{\circ}{M}{}_1^*$. Die Bildfolge $\overset{\circ}{\tau}(m_1^{(\nu)})$ hat sicher einen über $\mu_1(m_1) \in M$ liegenden Häufungspunkt $m_2 \in M_2^*$, da $\mu_1(m_1^{(\nu)})$ gegen $\mu_1(m_1)$ konvergiert und μ_2 eine eigentliche Abbildung ist. Wir behaupten, daß die Folge $\overset{\circ}{\tau}(m_1^{(\nu)})$ sogar gegen m_2 konvergiert.

Um das einzusehen, ordnen wir jedem Punkt $m_2^{(j)} \in \mu_2^{-1}(\mu_1(m_1))$ (es gibt nur endlich viele solche Punkte!) eine Umgebung $U_j(m_2^{(j)}) \subset M_2^*$ zu, derart, daß U_{j_1} und U_{j_2} für $j_1 \neq j_2$ punktfremd sind. Sind $\overset{\circ}{V}_j = \overset{\circ}{\tau}{}^{-1}(U_j)$ die Urbilder der U_j in $\overset{\circ}{M}{}_1^*$ bezüglich $\overset{\circ}{\tau}$, so gibt es sicher eine zusammenhängende Umgebung $V(m_1)$, derart, daß $\overset{\circ}{V} = V(m_1) \cap \overset{\circ}{M}{}_1^*$ in der offenen Menge $\bigcup_j \overset{\circ}{V}_j$ enthalten ist [anderenfalls gäbe es nämlich eine gegen m_1 konvergierende Punktfolge $\tilde{m}_1^{(\nu)} \in \overset{\circ}{M}{}_1^*$, deren Bildfolge $\overset{\circ}{\tau}(\tilde{m}_1^{(\nu)})$ sich nicht gegen einen über $\mu_1(m_1)$ liegenden Punkt von M_2^* häufen würde]. Da die Menge $M_1^* - \overset{\circ}{M}{}_1^*$ den Raum M_1 nirgends zerlegt, ist mit V auch die Menge $\overset{\circ}{V}$ zusammenhängend. Alsdann wird $\overset{\circ}{V}$ vermöge $\overset{\circ}{\tau}$ in eine bestimmte Menge U_{j_0} abgebildet. Das bedeutet aber, daß $\overset{\circ}{\tau}(m_1^{(\nu)})$ nur den Punkt $m_2^{(j_0)}$ als Häufungspunkt haben kann, denn dies ist der einzige Punkt aus $\mu_2^{-1}(\mu_1(m_1))$, der in U_{j_0} liegt. Mithin hat die Folge $\overset{\circ}{\tau}(m_1^{(\nu)})$ genau einen Häufungspunkt; da aber andererseits jede ihrer Teilfolgen einen Häufungspunkt hat, konvergiert $\overset{\circ}{\tau}(m_1^{(\nu)})$ gegen $m_2 = m_2^{(j_0)}$.

Wir setzen nun: $\tau(m_1) = \lim_{\nu \to \infty} \tau(m_1^{(\nu)}) = m_2$. Diese Definition ist auf Grund des soeben bewiesenen unabhängig von der Wahl der Folge $m_1^{(\nu)}$; also ist τ eindeutig definiert und offensichtlich die stetige Fortsetzung von $\overset{\circ}{\tau}$ auf M_1^*. Ebenso kann man $\overset{\circ}{\tau}{}^{-1}$ stetig auf M_2^* fortsetzen. Aus Stetigkeitsgründen ist diese Fortsetzung die Umkehrung von τ; daher ist τ eine topologische Abbildung von M_1^* auf M_2^*. Nach Definition gilt: $\mu_1 = \mu_2 \circ \tau$.

Wir haben noch zu zeigen, daß τ und τ^{-1} holomorphe Abbildungen sind. Es genügt, diesen Beweis für τ zu führen. Laut Definition ist τ genau dann holomorph, wenn folgendes gilt: Ist f irgendeine in einer offenen Menge $U \subset M_2^*$ holomorphe Funktion, so ist $f_1 = f \circ \tau$ in $W = \tau^{-1}(U)$ holomorph. Sicher ist die Funktion f_1 in $\overset{\circ}{W} = W \cap \overset{\circ}{M}{}_1^*$ holomorph, da dort gilt: $f_1 = f \circ \overset{\circ}{\tau}$ und $\overset{\circ}{\tau}$ eine holomorphe Abbildung ist. Nun ist $W - \overset{\circ}{W} = \mu_1^{-1}(N) \cap W$ eine analytische Menge in W, deren Dimension in jedem ihrer Punkte kleiner ist als die Dimension von W. Da f_1 in ganz W stetig ist, ergibt sich die Holomorphie von f_1 in W aus dem folgenden

Satz von RIEMANN: *Eine in einem komplexen Raum stetige, außerhalb einer dünnen Menge erster Ordnung holomorphe Funktion ist schlechthin holomorph*[18]).

3. Die Existenz komplexer Überlagerungsräume zu analytischen Mengen in komplexen Mannigfaltigkeiten war bereits — wenigstens im Prinzip — WEIERSTRASS bekannt[19]). In der Sprache der modernen Mathematik findet sich ein Beweis in den hektographierten Seminarberichten von CARTAN ([*13*], Exp. XIII—XIV, [*15*], Exp. VI—IX). Die Darstellung von CARTAN zeigt, daß die Abbildung μ genau dann eine topologische Abbildung von M^* auf M ist, wenn M *lokal-irreduzibel* ist. Ein Punkt $x \in M$ besitzt nämlich genau so viele verschiedene Urbilder in M^* wie M in x irreduzible Komponenten hat[20]). Die Konstruktion von CARTAN führt weiter genau dann zu einem zusammenhängenden komplexen Raum M^*, wenn die analytische Menge M irreduzibel in X ist.

Somit gilt in Verbindung mit dem Einzigkeitssatz:

Satz 2. *Ist M eine analytische Menge in einer komplexen Mannigfaltigkeit X, so besitzt M (bis auf analytische Isomorphie) genau einen komplexen Überlagerungsraum (M^*, μ). Es gelten folgende Aussagen:*

1) *Ein Punkt $x \in M$ hat genau dann k verschiedene Urbilder in M^*, wenn M in x in k irreduzible Komponenten zerfällt.*

2) *Der Raum M^* wird dann und nur dann durch μ topologisch auf M abgebildet, wenn M lokal-irreduzibel ist.*

3) *Der Raum M^* ist genau dann zusammenhängend, wenn M eine irreduzible analytische Menge in X ist.*

Aus diesem Satz folgt insbesondere:

Jede in einer komplexen Mannigfaltigkeit X lokal-irreduzible analytische Menge M besitzt eine natürliche komplex-analytische Struktur und ist somit ein komplexer Raum.

4. Wir geben in diesem Abschnitt ein Beispiel eines komplexen Raumes an, welches lehrt, daß gewisse Aussagen der komplexen Analysis, die für

[18]) Zum Beweise siehe [*31*], p. 186ff. sowie [*33*], p. 417.

[19]) Vgl. hierzu [*42*], p. 100.

[20]) Eine in einer komplexen Mannigfaltigkeit X analytische Menge M heißt *reduzibel* in X, wenn es zwei nicht leere, von M verschiedene, in X analytische Mengen M' und M'' gibt, deren Vereinigung M ist. Andernfalls heißt M *irreduzibel* in X.

Die in X analytische Menge M heißt *irreduzibel im Punkte* $x \in X$, wenn es beliebig kleine Umgebungen $U(x)$ gibt, derart, daß $M \cap U(x)$ irreduzibel in $U(x)$ ist.

Über die lokale Zerlegung einer analytischen Menge M gilt: *Es gibt zu jedem Punkt $x \in M$ eine Umgebung U, derart, daß $M \cap U$ sich darstellt als Vereinigung von endlich vielen in U irreduziblen analytischen Mengen M_σ, $\sigma = 1, \ldots, s$, mit folgenden Eigenschaften:*

a) *Es gibt kein s_0, derart, daß M_{s_0} in der Menge $\bigcup_{\sigma \neq s_0} M_\sigma$ enthalten ist.*

b) *Jede Menge M_σ enthält den Punkt x und ist in x irreduzibel.*

Die Mengen M_σ, $\sigma = 1, \ldots, s$, heißen die *irreduziblen Komponenten* von M in x; ihre Anzahl s ist unabhängig von der Wahl der Umgebung U.

Die in X analytische Menge M heißt *lokal-irreduzibel*, wenn sie in jedem ihrer Punkte irreduzibel ist.

den C^n und allgemein für komplexe Mannigfaltigkeiten richtig sind, in komplexen Räumen ihre Allgemeingültigkeit verlieren.

Es seien $C^2_\varkappa$, $\varkappa = 1, \ldots, k$, $k \geq 2$, k Exemplare des 2-dimensionalen komplexen Zahlenraumes der Veränderlichen $z_\varkappa$, $w_\varkappa$. Im kartesischen Produkt $C^{2k} = \bigtimes_{\varkappa=1}^{k} C^2_\varkappa$ werde die analytische Menge

$$Y: \quad \left\{\frac{z_1}{w_1} = \cdots = \frac{z_k}{w_k}\right\}$$

betrachtet[21]). Alle vom Nullpunkt $O \in C^{2k}$ verschiedenen Punkte $y \in Y$ sind gewöhnliche Punkte von Y, da man das angeschriebene Gleichungssystem in der Umgebung eines jeden solchen Punktes immer nach $(k-1)$ Variablen auflösen kann. Daraus folgt insbesondere, daß Y in jedem Punkt $y \neq O$ die Dimension $k+1$ hat; das gilt alsdann auch noch für den Nullpunkt O selbst, da er Häufungspunkt von gewöhnlichen Punkten ist (vgl. etwa [35], p. 282). Mithin ist Y eine *rein $(k+1)$-dimensionale analytische Menge im C^{2k}*; wir schreiben daher auch: $Y = Y^{k+1}$.

Die Menge Y^{k+1} wird durch die k-dimensionalen analytischen Ebenen

$$E(s): \quad \left\{\frac{z_1}{w_1} = \cdots = \frac{z_k}{w_k} = s\right\}$$

vollständig überdeckt, wenn s die Zahlen einer RIEMANNschen Zahlenkugel P^1 durchläuft: $Y^{k+1} = \{E(s), s \in P^1\}$. Zwei Ebenen $E(s_1)$ und $E(s_2)$, $s_1 \neq s_2$, schneiden sich genau im Punkte $O \in C^{2k}$.

Wir behaupten nun: *Y ist im Nullpunkt O irreduzibel.*

Wir zeigen dazu, daß in keiner Kugel K um O die Menge $Y \cap K$ in zwei nicht leere, von $Y \cap K$ verschiedene analytische Mengen Y_1, Y_2 zerfallen kann. Könnte man nämlich für eine gewisse Kugel K die Menge $Y \cap K$ als Vereinigung zweier solcher, dann notwendig rein $(k+1)$-dimensionaler analytischer Mengen Y_1 und Y_2 darstellen, so würden Y_1 und Y_2 mit eventueller Ausnahme von O nur gewöhnliche Punkte besitzen. Daher könnte $Y \cap (K - O)$ nicht zusammenhängend sein. Das ist aber der Fall, denn offensichtlich ist die Menge $\{E(s) \cap (K - O), s \in P^1\}$, die mit $Y \cap (K - O)$ übereinstimmt, eine zusammenhängende Punktmenge. Y^{k+1} ist somit eine lokal-irreduzible, rein $(k+1)$-dimensionale analytische Menge im C^{2k} und besitzt folglich nach § 2.3 eine natürliche komplex-analytische Struktur. Wir können daher Y^{k+1} als einen $(k+1)$-dimensionalen komplexen Raum auffassen; offensichtlich ist Y^{k+1} zusammenhängend.

5. Wir bezeichnen im folgenden den Nullpunkt des C^{2k} mit y_0 und zeigen, daß $y_0 \in Y^{k+1}$ ein nicht-uniformisierbarer Punkt von Y^{k+1} ist. Zu dem Zwecke

[21]) Werden die Koordinaten des C^{2k} als homogene Koordinaten des $(2k-1)$-dimensionalen projektiven Raumes P^{2k-1} gedeutet, so geht der Kegel Y in eine k-dimensionale algebraische Menge des P^{2k-1} über, die in der algebraischen Geometrie seit langem bekannt ist und dort als Segremannigfaltigkeit $S_{k-1,1}$ bezeichnet wird. Y selbst würde man am besten Segrekegel nennen. Vgl. dazu etwa BURAU, W.: Grundmannigfaltigkeiten der projektiven Geometrie. Collectanea Math. **1950–1953**.

weisen wir nach, daß in einer Umgebung von y_0 nicht mehr alle Eigenschaften gelten, die von der komplexen Analysis auf komplexen Mannigfaltigkeiten her bekannt sind. Für komplexe Mannigfaltigkeiten gilt z.B. der folgende Satz (vgl. [37]):

Sind M_1 und M_2 zwei rein d_1- bzw. d_2-dimensionale analytische Mengen in einer rein n-dimensionalen komplexen Mannigfaltigkeit X, so ist $M_1 \cap M_2$ leer oder eine mindestens (d_1+d_2-n)-dimensionale analytische Menge in X.

Für den hier konstruierten komplexen Raum verliert dieser Satz aber schon seine Gültigkeit im Falle $d_1=d_2=n-1$. Es gilt nämlich, wenn man $n=k+1$ setzt (man beachte $n \geq 3$):

Satz 3. *Die im rein n-dimensionalen komplexen Raum Y^n rein $(n-1)$-dimensionalen analytischen Mengen $E(s_1)$, $E(s_2)$, $s_1 \neq s_2$, schneiden sich nur im Punkte y_0, also in einer nicht leeren analytischen Menge, deren Dimension kleiner als $n-2$ ist.*

Für komplexe Mannigfaltigkeiten gilt weiter (vgl. [35], p. 276):

Jede in einer n-dimensionalen komplexen Mannigfaltigkeit X rein $(n-1)$-dimensionale analytische Menge ist lokal stets das Nullstellengebilde genau einer holomorphen Funktion.

Auch dieser Satz wird im Raum Y^n falsch; vielmehr läßt sich zeigen:

Satz 4. *Die in Y^n rein $(n-1)$-dimensionale analytische Menge $E(s)$ ist in einer Umgebung von y_0 nicht als simultanes Nullstellengebilde von weniger als $(n-1)$ lokal-holomorphen Funktionen darstellbar.*

In der Tat: Ließe sich nämlich die Ebene $E(s)$ in einer Umgebung $U \subset Y^n$ des Punktes y_0 als simultanes Nullstellengebilde von weniger als $(n-1)$ in U holomorphen Funktionen $g_1, \ldots, g_m$ darstellen, so würden die Funktionen $g_1, \ldots, g_m$ auf jeder von $E(s)$ verschiedenen Ebene $E(s')$ in $U \cap E(s')$ nur im Nullpunkt gemeinsam verschwinden. Das ist aber nicht möglich, denn aus dem oben für komplexe Mannigfaltigkeiten angegebenen Satz ergibt sich durch einen einfachen Induktionsschluß, wenn man beachtet, daß $E(s')$ eine rein $(n-1)$-dimensionale komplexe Mannigfaltigkeit ist, daß die analytische Menge

$$\{g_1=0, g_2=0, \ldots, g_m=0\}$$

in $U \cap E(s')$ mindestens $(n-m-1)$-dimensional ist. Wegen $m<n-1$ haben also die $g_1, \ldots, g_m$ noch weitere gemeinsame Nullstellen in $U \cap E(s')$.

Die Aussage von Satz 4 ist besonders überraschend, da sich andererseits zeigen läßt:

Ist X ein n-dimensionaler komplexer Raum, so läßt sich jede analytische Menge in X lokal stets als simultanes Nullstellengebilde von höchstens n lokal-holomorphen Funktionen darstellen[22]).

Weiter zeigt der Raum Y^n bezüglich meromorpher Funktionen ein anderes Verhalten als komplexe Mannigfaltigkeiten. Für rein n-dimensionale komplexe Mannigfaltigkeiten X gilt bekanntlich:

[22]) Vgl. dazu [20], Satz 1. Mit dem dort angegebenen Verfahren läßt sich diese Aussage beweisen.

Die außerwesentlichen Singularitäten zweiter Art (Unbestimmtheitsstellen) einer in X meromorphen Funktion bilden eine leere oder rein $(n-2)$-dimensionale analytische Menge in X[23]).

Für den Raum Y^n gilt dagegen:

Satz 5. *Es gibt eine in Y^n meromorphe Funktion $h(y)$, die als einzige Unbestimmtheitsstelle den Punkt y_0 hat.*

Man hat nur zu setzen: $h(y) = \frac{z_1}{w_1} = \cdots = \frac{z_k}{w_k} = s$.

6. Der Raum Y^n läßt sich als endlich-blättrige, analytisch-verzweigte Überlagerung des C^n auffassen. Um das einzusehen, definieren wir zunächst eine holomorphe Abbildung Φ von Y^n auf den Raum C^n, $n = k+1$, der Veränderlichen $v_0, v_1, \ldots, v_k$, indem wir setzen:

$$\Phi: \begin{cases} v_0 = z_1 + \cdots + z_k, \\ v_1 = w_1 + a_1 z_1, \\ \vdots \\ v_k = w_k + a_k z_k, \end{cases} \qquad (z_1, w_1, \ldots, z_k, w_k) \in Y^n;$$

dabei seien $a_1, \ldots, a_k$ fest gewählte, paarweise voneinander verschiedene komplexe Zahlen vom Betrage 1.

Offensichtlich hat der Nullpunkt $O \in C^n$ nur den Punkt $y_0 \in Y^n$ als Φ-Urbild. Durch Φ wird ferner, wie man sofort nachrechnet, die $(n-1)$-dimensionale analytische Ebene $E(s) < Y^n$ linear auf die $(n-1)$-dimensionale analytische Ebene

$$v_0 \cdot \prod_{\varkappa=1}^{k} (1 + a_\varkappa s) - s \sum_{\varkappa=1}^{k} v_\varkappa \prod_{\substack{\nu=1 \\ \nu \neq \varkappa}}^{k} (1 + a_\nu s) = 0$$

im C^n abgebildet. Daraus folgt, daß jeder Punkt $\mathfrak{v} = (v_0, v_1, \ldots, v_k)$ des C^n in jeder Ebene $E(s)$ höchstens einen Urbildpunkt besitzt. Die Parameter s der Ebenen $E(s)$, in denen ein solcher Urbildpunkt liegt, müssen notwendig das Polynom in s: $P(s; \mathfrak{v}) = v_0 \cdot \prod_{\varkappa=1}^{k} (1 + a_\varkappa s) - s \cdot \sum_{\varkappa=1}^{k} v_\varkappa \cdot \prod_{\substack{\nu=1 \\ \nu \neq \varkappa}}^{k} (1 + a_\nu s)$ annullieren, oder es muß $s = \infty$ sein. Da $P(s; \mathfrak{v})$ nur für $\mathfrak{v} = 0$ identisch verschwindet und im übrigen bezüglich s stets höchstens vom Grade k ist, besitzt also jeder Punkt $\mathfrak{v} \in C^n$ höchstens k verschiedene Urbilder in Y^n; die Abbildung Φ ist daher eine nirgends entartete Abbildung. Es folgt auch: Φ ist eine eigentliche Abbildung von Y^n auf C^n.

Es gibt Punkte $\mathfrak{v} \in C^n$, derart, daß das Polynom $P(s; \mathfrak{v})$ genau k verschiedene Nullstellen hat, denn von den $(k+1)$ Koeffizienten von $P(s; \mathfrak{v})$ können k beliebig gewählt werden, da für $\mathfrak{v}$ alle Punkte des C^n zulässig sind. Gewisse Punkte des C^n besitzen also genau k verschiedene Urbilder in Y^n.

Nach diesen Bemerkungen ergibt sich leicht:

Satz 6. *Das Tripel (Y^n, Φ, C^n) ist eine $(n-1)$-blättrige, analytisch-verzweigte Überlagerung des C^n.*

[23]) Vgl. hierzu [2], p. 61, sowie [31], p. 182.

Beweis. Die Bedingung 0) von Def. 1 ist bereits verifiziert. Um Bedingung 1) zu bestätigen, betrachten wir in der komplexen Mannigfaltigkeit $Y^n - y_0$ die Nullstellenmenge A' der Funktionaldeterminante der Abbildung Φ. A' ist eine rein $(n-1)$-dimensionale analytische Menge in $Y^n - y_0$. Da Φ eine nirgends entartete eigentliche holomorphe Abbildung von $Y^n - y_0$ auf $C^n - O$ ist, wird A' nach einem für komplexe Räume geltenden Abbildungssatz (vgl. [33], [34]) durch Φ auf eine rein $(n-1)$-dimensionale analytische Menge A^* in $C^n - O$ abgebildet. A^* ist nach einem Fortsetzungssatz über analytische Mengen (vgl. [40], [35]) zu einer im C^n analytischen Menge A fortsetzbar, wir dürfen $O \in A$ annehmen. Die Menge $\Phi^{-1}(A)$ ist eine analytische Menge in Y^n und zerlegt daher Y^n nirgends. Da $\Phi^{-1}(A)$ die Menge $A' \cup y_0$ umfaßt, ist Φ in $Y^n - \Phi^{-1}(A)$ lokal-topologisch. Damit ist aber Bedingung 1) als erfüllt nachgewiesen, w. z. b. w.

7. Der in dem $(2n-2)$-dimensionalen komplexen Zahlenraum C^{2n-2} holomorph eingebettete n-dimensionale komplexe Raum Y^n hat weiter die folgende bemerkenswerte Eigenschaft:

Satz 7. *Es gibt keine Umgebung $U \subset Y^n$ des Punktes $y_0 \in Y^n$, die umkehrbar holomorph auf eine in einem Gebiet $G \subset C^r$, $r < 2n-2$ lokal-irreduzible analytische Menge abbildbar ist.*

Andernfalls hätten nämlich zwei verschiedene Ebenen $E(s_1)$ und $E(s_2)$ eine mindestens eindimensionale analytische Menge miteinander gemeinsam[24]).

Der n-dimensionale komplexe Raum Y^n kann also sicher nicht dadurch erzeugt werden, daß man eine in einem Zahlenraum C^r mit $r < 2n-2$ liegende lokal-irreduzible analytische Menge mit der natürlichen komplex-analytischen Struktur versieht. Jedoch gibt es bereits im C^{n+1} analytische Mengen, die Y^n als Überlagerungsraum haben.

Satz 8. *Es gibt im $(n+1)$-dimensionalen komplexen Zahlenraum C^{n+1} eine analytische Menge M und eine holomorphe Abbildung μ von Y^n auf M, derart, daß (Y^n, μ) der komplexe Überlagerungsraum von M ist.*

Beweis. Sei $v^* \in C^n$ ein Punkt, über dem kein Verzweigungspunkt der analytisch-verzweigten Überlagerung (Y^n, Φ, C^n) liegt. Sei f eine holomorphe Funktion auf Y^n, die in den über v^* liegenden Punkten von Y^n jeweils verschiedene Werte hat (man wähle etwa für f ein lineares homogenes Polynom in den $z_1, w_1, z_2, w_2, \ldots, z_k, w_k$). Die Funktion f annulliert nach bekannten Sätzen[25]) ein irreduzibles Pseudopolynom

$$\tilde{\omega}(v; v_0, v_1, \ldots, v_k) \equiv v^k + c_1(v_0, v_1, \ldots, v_k)\, v^{k-1} + \cdots + c_k(v_0, v_1, \ldots, v_k).$$

Die Diskriminantenmenge D von $\tilde{\omega}$ ist eine den Nullpunkt des C^n umfassende analytische Menge und läßt sich als Nullstellengebilde einer holomorphen Funktion $h(v_0, \ldots, v_k)$ darstellen. Wir betrachten alsdann auf Y^n

[24]) Man muß benutzen, daß die Bilder der Ebenen $E(s_1)$ und $E(s_2)$ rein $(n-1)$-dimensionale analytische Mengen in G sind. Das aber folgt aus einem in [33], [34] bewiesenen Projektionssatz.

[25]) Zum Begriff des Pseudopolynoms sowie der Diskriminante vgl. [31].

das Produkt g der Funktionen f und $h \circ \Phi$. Die Funktion g ist ebenfalls holomorph auf Y^n und zwar die Nullstelle des irreduziblen Pseudopolynoms

$$\omega(v; v_0, \ldots, v_k) \equiv v^k + c_1(v_0, \ldots, v_k) \cdot h(v_0, \ldots, v_k)\, v^{k-1} + \cdots$$
$$\cdots + c_k(v_0, \ldots, v_k) \cdot h(v_0, \ldots, v_k)^k.$$

Die Pseudopolynome $\tilde{\omega}$ und ω haben offenbar dieselbe Diskriminantenmenge D. Bezeichnet nun M die im Raum C^{n+1} der Veränderlichen $v, v_0, v_1, \ldots, v_k$ gelegene Nullstellenmenge des Pseudopolynoms ω, so besitzt M sicher keine gewöhnlichen Punkte in der Menge $\{v\} \times D$. Denn alle partiellen Ableitungen von ω verschwinden dort, da die Koeffizienten von ω sämtlich in D verschwinden, und zwar das konstante Glied sogar in mindestens k-ter Ordnung ($k \geq 2$). Da andererseits jeder Punkt von M, der nicht zu $\{v\} \times D$ gehört, sicher ein gewöhnlicher Punkt von M ist, so besteht die Menge N der nichtgewöhnlichen Punkte von M also genau aus der Menge $M \cap (\{v\} \times D)$.

Wir betrachten nun die holomorphe Abbildung $\mu = g \times \Phi$ von Y^n in den C^{n+1}. Dieselbe bildet Y^n auf M ab und ist offensichtlich nirgends entartet und eigentlich. Nach dem Vorstehenden gilt: $\mu^{-1}(N) = \Phi^{-1}(D)$. Da g in allen Punkten von Y^n, die über einem Punkt $\mathfrak{v} \notin D$ liegen, paarweise verschiedene Werte besitzt, wird also durch μ die Menge $Y^n - \Phi^{-1}(D)$ eineindeutig auf die Menge der gewöhnlichen Punkte von M abgebildet. Da $Y^n - \Phi^{-1}(D)$ eine komplexe Mannigfaltigkeit ist [es gilt $y_0 \in \Phi^{-1}(D)$], so ist $\mu: Y^n - \Phi^{-1}(D) \to M - N$ nach einem bekannten Satze[26]) holomorph umkehrbar. Damit ist alles bewiesen.

Wir können die Menge $M \subset C^{n+1}$ noch näher beschreiben:

Satz 9. *Die im C^{n+1} analytische Menge M ist irreduzibel und n-dimensional. M ist im Nullpunkt irreduzibel, jedoch gibt es in beliebiger Nähe des Nullpunktes Punkte von M, in denen M reduzibel ist*[27]).

Der Beweis ist einfach. Da Y^n ein zusammenhängender n-dimensionaler komplexer Raum ist, so ist M wegen Satz 2 irreduzibel und n-dimensional. Da über $O \in M$ nur der Punkt $y_0 \in Y^n$ liegt, folgt ebenfalls aus Satz 2, daß M in O irreduzibel ist. Es muß aber in beliebiger Nähe von O Punkte von M geben, in denen M reduzibel ist, da sonst nach Satz 2 die Abbildung μ eine Umgebung U von $y_0 \in Y^n$ umkehrbar holomorph auf eine Umgebung $V(O) \in M$ abbilden würde, was nach Satz 7 nicht möglich ist.

Der Menge M läßt sich auf Grund von Satz 9 in keiner Umgebung $V(O) \subset M$ eine komplex-analytische Struktur aufprägen, derart, daß die Injektion $\iota: M \to C^{n+1}$ eine holomorphe Abbildung ist. Denn wäre das möglich, so müßte $\iota^{-1}(N)$, wo N die Menge der nichtgewöhnlichen Punkte von M ist, eine höchstens $(n-1)$-dimensionale analytische Menge im komplexen Raum M sein, die M alsdann lokal nicht zerlegen könnte. Nun ist aber M in Punkten $\mathfrak{v}'$,

[26]) Nach [31], p. 149, sind eineindeutige holomorphe Abbildungen gleichdimensionaler komplexer Mannigfaltigkeiten aufeinander holomorph umkehrbar.

[2]) Es gibt bereits im C^3 zweidimensionale analytische Mengen, die im Nullpunkt O irreduzibel, jedoch in keiner Umgebung von O lokal-irreduzibel sind. Das obige Beispiel gilt dagegen nur für $n \geq 3$.

die beliebig nahe bei O gewählt werden können, reduzibel; in der Umgebung eines jeden solchen Punktes $\mathfrak{v}'$ zerfällt also M in mindestens zwei irreduzible Komponenten M_1 und M_2. Da offensichtlich $M_1 - N$ und $M_2 - N$ disjunkt sind, so zerlegt also N den Raum M in jeder Umgebung von $\mathfrak{v}'$.

§ 3. RIEMANNsche Gebiete

1. Bereits im Jahre 1932 führten Untersuchungen von CARTAN und THULLEN zu der Einsicht, daß man in der Funktionentheorie mehrerer Veränderlichen neben den schlichten Gebieten im C^n zugleich auch nichtschlichte Gebiete über dem C^n betrachten muß (vgl. [39], p. 76 sowie [9], [10]). Im Laufe der Zeit sind für diese sog. RIEMANNschen Gebiete verschiedene Definitionen angegeben worden, die jedoch sämtlich die Verzweigungspunkte unberücksichtigt lassen. Wir werden in diesem Paragraphen verzweigte RIEMANNsche Gebiete als spezielle komplexe Räume definieren, die in bestimmter Art dem C^n überlagert sind.

Def. 16 (*RIEMANNsches Gebiet*). *Ein Paar $\mathfrak{G} = (G, \Phi)$ heißt ein RIEMANNsches Gebiet über dem C^n, wenn die folgenden beiden Bedingungen erfüllt sind:*

0) *G ist ein HAUSDORFFscher Raum und Φ eine Abbildung von G in den C^n.*

1) *Zu jedem Punkt $x \in G$ gibt es Umgebungen $U(x) \subset G$, $V(\Phi(x)) \subset C^n$, so daß (U, Φ, V) eine endlich-blättrige, analytisch-verzweigte Überlagerung ist.*

Aus 1) folgt sofort, daß Φ eine stetige, nirgends entartete Abbildung von G in den C^n ist. Ferner wird G durch 1) in natürlicher Weise mit einer komplexen Struktur versehen. G kann deshalb als komplexer Raum aufgefaßt werden; die Abbildung $\Phi: G \to C^n$ ist holomorph.

Def. 17 (*Unverzweigtes RIEMANNsches Gebiet*). *Ein RIEMANNsches Gebiet $\mathfrak{G} = (G, \Phi)$ heißt unverzweigt, wenn Φ eine lokal-topologische Abbildung ist.*

Für gewisse weitere Überlegungen benötigen wir den Begriff des (erreichbaren) Randpunktes.

Def. 18 (*Randpunkt*)[28]. *Ein (erreichbarer) Randpunkt eines RIEMANNschen Gebietes $\mathfrak{G} = (G, \Phi)$ ist ein Filter*[29] *$\mathfrak{r}$ von offenen, zusammenhängenden Mengen $U \subset G$ mit folgenden Eigenschaften:*

1) *$\mathfrak{r}$ hat in G keinen Häufungspunkt;*

2) *der Bildfilter $\{\Phi(U), U \in \mathfrak{r}\}$ konvergiert gegen einen Punkt $\mathfrak{z}_0 \in C^n$;*

3) *ist V eine Umgebung von $\mathfrak{z}_0$, so enthält $\mathfrak{r}$ genau eine zusammenhängende Komponente von $\Phi^{-1}(V)$. Alle Mengen $U \in \mathfrak{r}$ werden auf diese Weise erzeugt.*

Die Menge aller (erreichbaren) Randpunkte von $\mathfrak{G}$ heißt der *Rand* ∂G von $\mathfrak{G}$. Die Abbildung Φ wird durch die Setzung $\breve{\Phi}(\mathfrak{r}) = \mathfrak{z}_0$, $\breve{\Phi}(x) = \Phi(x)$ für $\mathfrak{r} \in \partial G$, $x \in G$ zu einer Abbildung $\breve{\Phi}$ von $\breve{G} = G \cup \partial G$ in den C^n fortgesetzt. Wir führen in $\breve{G}$ eine Topologie ein, indem wir den Randpunkten $\mathfrak{r}_0 \in \partial G$

[28]) Eine äquivalente Definition des Randpunktes findet sich in [2], p. 13.

[29]) Unter einem *Filter* $\mathfrak{F}$ auf einer nichtleeren Menge M verstehen wir ein nichtleeres Mengensystem nichtleerer Mengen, derart, daß zu zwei Mengen $N_1, N_2 \in \mathfrak{F}$ stets eine Menge $N_3 \in \mathfrak{F}$ existiert, die in $N_1 \cap N_2$ enthalten ist.

Umgebungen zuordnen. Es sei etwa U_0 eine Menge aus r_0. Unter einer Umgebung $V(r_0)$ verstehen wir dann die Vereinigung von U_0 mit den Filtern $r \in \partial G$, die wenigstens eine Menge $U \subset U_0$ enthalten.

Offenbar wird $\breve{G}$ durch diese Topologisierung zu einem HAUSDORFFschen Raum; $\breve{\Phi}$ ist eine stetige Abbildung von $\breve{G}$ in den C^n.

Wir merken noch an:

Zu jedem Punkt $r \in \partial G$ gibt es eine Kurve $\{x(t), 0 \leq t \leq 1\}$, $x(1) = r$, in $\breve{G}$, die für $0 \leq t < 1$ in G verläuft.

Ist umgekehrt $\{x(t), 0 \leq t < 1\}$ eine Kurve in G und läßt sich die Abbildung $\Phi \circ x(t)$ zu einer stetigen Abbildung von $0 \leq t \leq 1$ in den C^n fortsetzen, so kann man in $\breve{G}$ auch $\{x(t), 0 \leq t < 1\}$ in eindeutiger Weise zu einer Kurve $\{x(t), 0 \leq t \leq 1\}$ fortsetzen. — Der Beweis beider Aussagen liegt auf der Hand.

2. Die RIEMANNschen Gebiete über dem C^n, $n \geq 2$, bilden eine zu umfangreiche Klasse von komplexen Räumen, als daß sich in ihnen im allgemeinen eine inhaltsvolle Funktionentheorie entwickeln ließe. Dementsprechend ist die Menge der analytischen Gebilde holomorpher Funktionen auch wesentlich kleiner als die Klasse der RIEMANNschen Gebiete. Es empfiehlt sich deshalb, alle funktionentheoretischen Untersuchungen auf die Existenzgebiete holomorpher Funktionen, die sog. Holomorphiegebiete, zu beschränken. Um diesen Begriff zu präzisieren, geben wir zunächst folgende Definitionen:

Def. 19 *(Enthaltensein). Ein* RIEMANNsches *Gebiet $\mathfrak{G} = (G, \Phi)$ ist in einem* RIEMANNschen *Gebiet $\mathfrak{G}' = (G', \Phi')$ enthalten, wenn es eine spurpunkttreue, stetige Abbildung τ von G in G' gibt.*

Def. 20 *(Äquivalenz). Zwei* RIEMANNsche *Gebiete $\mathfrak{G} = (G, \Phi)$, $\mathfrak{G}' = (G', \Phi')$ heißen äquivalent (in Zeichen $\mathfrak{G} = \mathfrak{G}'$), wenn sich G topologisch und spurpunkttreu auf G' abbilden läßt.*

Def. 21 *(Fortsetzung). Es sei $\mathfrak{G} = (G, \Phi)$ ein* RIEMANNsches *Gebiet, das vermöge einer stetigen, spurpunkttreuen Abbildung $\tau: G \to G'$ in einem* RIEMANNschen *Gebiet $\mathfrak{G}' = (G', \Phi')$ enthalten ist. Ist dann $f(x)$ eine in G und $f'(x')$ eine in G' holomorphe Funktion, so heißt $f'(x')$ die holomorphe Fortsetzung von $f(x)$ in $\mathfrak{G}'$, wenn in G gilt: $f(x) = f' \circ \tau(x)$.*

Wir verwenden nun diese Definitionen, um die Holomorphiegebiete einzuführen:

Def. 22 *(Holomorphiegebiet). Ein* RIEMANNsches *Gebiet $\mathfrak{G} = (G, \Phi)$ heißt ein Holomorphiegebiet, wenn es eine in G holomorphe Funktion f gibt, die sich in kein $\mathfrak{G}$ umfassendes, zu $\mathfrak{G}$ nicht äquivalentes* RIEMANNsches *Gebiet fortsetzen läßt.*

Eine spezielle Klasse von Holomorphiegebieten bilden die holomorph-konvexen RIEMANNschen Gebiete. Der hier auftretende Begriff der Holomorphiekonvexität läßt sich in beliebigen komplexen Räumen X einführen. Wir definieren dazu zunächst die holomorph-konvexe Hülle einer beliebigen Teilmenge von X (vgl. hierzu etwa auch [14]):

Def. 23 *(Holomorph-konvexe Hülle). Unter der holomorph-konvexen Hülle* $\hat{M}$ *einer Teilmenge M eines komplexen Raumes X versteht man die Menge* $\{x \in X, |f(x)| \leq \sup |f(M)|\}$, *wobei f alle in X holomorphen Funktionen durchläuft.*

Offenbar ist $\hat{M}$ immer abgeschlossen. Wir definieren weiter:

Def. 24 *(Holomorphiekonvexität). Ein komplexer Raum heißt holomorph-konvex, wenn die holomorph-konvexe Hülle jeder kompakten Teilmenge von X kompakt ist.*

Ein RIEMANNsches *Gebiet* $\mathfrak{G} = (G, \Phi)$ *heißt holomorph-konvex, wenn der komplexe Raum G holomorph-konvex ist.*

Es gilt nun:

Satz 10. *Jedes holomorph-konvexe* RIEMANNsche *Gebiet, dessen Verzweigungsverhalten lokal jeweils durch eine holomorphe Funktion repräsentiert werden kann, ist ein Holomorphiegebiet*[30]).

OKA zeigte 1953, daß die unverzweigten Holomorphiegebiete durch die Holomorphiekonvexität charakterisiert werden. Er bewies nämlich [*30*]:

Satz 11. *Ein unverzweigtes* RIEMANNsches *Gebiet über dem* C^n *ist dann und nur dann ein Holomorphiegebiet, wenn es holomorph-konvex ist.*

Spezialfälle dieses sehr tiefliegenden Resultates waren seit längerem bekannt. So bewies bereits 1932 THULLEN [*39*] die Richtigkeit von Satz 11 für schlichte Holomorphiegebiete. Bei der Wichtigkeit des Problems übertrugen BEHNKE und STEIN die Aussage des THULLENschen Satzes auf die sog. fastendlich-blättrigen RIEMANNschen Gebiete [*4*]. Um jedoch den Allgemeinfall zu lösen, mußten von OKA völlig andere Methoden herausgearbeitet werden.

3. Als entscheidend für die Charakterisierung der Holomorphiegebiete durch lokale Eigenschaften (die Holomorphiekonvexität muß als globale Eigenschaft angesehen werden!), hat sich die Pseudokonvexität herausgestellt. Wir ziehen zur Definition spezielle Scharen analytischer Mengen heran (vgl. hierzu auch [*22*]):

Def. 25 *(Schar analytischer Mengen). Unter einer Schar eindimensionaler analytischer Mengen in einem* RIEMANNschen *Gebiet* $\mathfrak{G} = (G, \Phi)$ *wird eine nicht konstante, holomorphe Abbildung* $\sigma(w, t)$ *des Einheitskreises* $K: \{|w| \leq 1\}$ *in* $\breve{G}$ *verstanden, die noch stetig von einem reellen Parameter t,* $0 \leq t \leq 1$, *abhängt.*

Dabei sagen wir, $\sigma(w, t)$, $t = t_0$, bildet K holomorph in $\breve{G}$ ab, wenn σ stetig und die Abbildung $\breve{\Phi} \circ \sigma(w, t): K \to C^n$ in einer Umgebung von K holomorph ist. Wir setzen $\sigma(w, t) = \sigma_\varepsilon(w, t)$, wenn die Menge $\{\Phi \circ \sigma(w, t), |w| \leq 1, 0 \leq t \leq 1\}$ in einer Hyperkugel vom Radius ε enthalten ist.

Def. 26 *(Ausgezeichnete Schar). Eine Schar* $\sigma(w, t)$ *eindimensionaler analytischer Mengen in einem* RIEMANNschen *Gebiet* $\mathfrak{G} = (G, \Phi)$ *heißt aus-*

[30]) Ein Beweis dieses Satzes ergibt sich unmittelbar aus [*19*], Satz B, in Verbindung mit [*2*], Satz 37. Vgl. auch [*1*]. Am Schluß dieser Arbeit findet sich eine Zusammenstellung der wichtigsten Eigenschaften holomorph-konvexer RIEMANNscher Gebiete.

gezeichnet, wenn $\sigma(w, t)$ für $w \in K$, $0 \leq t < 1$, und für $w \in \partial K$, $0 \leq t \leq 1$ in G enthalten ist.

Es läßt sich nun definieren:

Def. 27 *(Pseudokonvexes* RIEMANNsches *Gebiet). Ein* RIEMANNsches *Gebiet $\mathfrak{G} = (G, \Phi)$ heißt pseudokonvex, wenn für jede in $\mathfrak{G}$ ausgezeichnete Schar eindimensionaler analytischer Mengen $\sigma(w, t)$ der Durchschnitt $\{\sigma(w, t)\} \cap \partial G$ leer ist.*

Def. 28 *(Pseudokonvexer Randpunkt). Ein Randpunkt r eines* RIEMANNschen *Gebietes $\mathfrak{G} = (G, \Phi)$ heißt pseudokonvex, wenn es eine Umgebung $U(r) \subset \breve{G}$ und ein $\varepsilon > 0$ gibt, so daß jede in $\mathfrak{G}$ ausgezeichnete Schar analytischer Mengen $\sigma_\varepsilon(w, t)$ mit $U(r) \cap \partial G$ keinen Punkt gemeinsam hat.*

Offenbar bilden die pseudokonvexen Randpunkte eine offene Menge in ∂G. Ist $\mathfrak{G}$ pseudokonvex, so ist auch jeder Randpunkt von $\mathfrak{G}$ pseudokonvex. Während aber die Pseudokonvexität selbst noch eine globale Eigenschaft ist, ist die Pseudokonvexität eines Randpunktes sicher rein lokaler Natur. OKA konnte nun zeigen [*30*]:

Satz 12. *Ein unverzweigtes* RIEMANNsches *Gebiet $\mathfrak{G}$ ist genau dann pseudokonvex, wenn jeder Randpunkt von $\mathfrak{G}$ pseudokonvex ist.*

Ferner ergab sich das sehr tiefliegende Resultat:

Satz 13. *Ein unverzweigtes* RIEMANNsches *Gebiet ist genau dann holomorph-konvex, wenn es pseudokonvex ist.*

Zusammen mit Satz 11 und Satz 12 folgt daraus:

Satz 14. *Ein unverzweigtes* RIEMANNsches *Gebiet $\mathfrak{G}$ ist dann und nur dann ein Holomorphiegebiet, wenn jeder Randpunkt von $\mathfrak{G}$ pseudokonvex ist.*

Es sei angemerkt, daß OKA diese Sätze nicht in der hier angegebenen Reihenfolge bewiesen hat. Er zeigte zunächst, daß jedes unverzweigte Holomorphiegebiet pseudokonvex ist und schloß erst dann, daß jedes unverzweigte pseudokonvexe RIEMANNsche Gebiet ein holomorph-konvexes Holomorphiegebiet ist.

4. Grundlegende Sätze der Funktionentheorie auf offenen RIEMANNschen Flächen gelten mutatis mutandis auch in unverzweigten Holomorphiegebieten. Der klassische Satz von MITTAG-LEFFLER sagt aus, daß man auf nicht-kompakten RIEMANNschen Flächen zu sinnvoll vorgegebenen Hauptteilen immer eine meromorphe Funktion konstruieren kann, die diese Hauptteile besitzt (vgl. etwa [*6*], p. 565 ff.). In der Funktionentheorie mehrerer Veränderlichen, wo die Polstellen meromorpher Funktionen analytische Mengen der Dimension $(n-1)$ bilden, läuft die Vorgabe von Hauptteilen auf die Vorgabe von sog. Cousin-I-Verteilungen hinaus [*16*]:

Def. 29 *(Cousin-I-Verteilung). Es sei X ein komplexer Raum; $\{U_\iota, \iota \in I\}$ sei eine offene Überdeckung von X. Eine Menge $\{f_\iota\}$ von in U_ι meromorphen Funktionen f_ι heißt dann eine Cousin-I-Verteilung in X, wenn in allen nicht leeren Durchschnitten $U_{\iota_1} \cap U_{\iota_2}$ die Differenz $f_{\iota_1} - f_{\iota_2}$ holomorph ist.*

Def. 30 *(Lösung einer Cousin-I-Verteilung). Eine in X meromorphe Funktion f heißt eine Lösung einer Cousin-I-Verteilung $\{f_\iota\}$, wenn jede Differenz $f-f_\iota$ in der offenen Menge U_ι holomorph ist.*

Es gilt nun als Verallgemeinerung des MITTAG-LEFFLERschen Satzes:

Satz 15. *In einem unverzweigten Holomorphiegebiet besitzt jede Cousin-I-Verteilung eine Lösung*[31]).

Für nicht-kompakte RIEMANNsche Flächen gilt ferner der Satz von WEIERSTRASS, der aussagt, daß man zu jeder sich auf der Fläche nicht häufenden Punktfolge x_ν eine holomorphe Funktion finden kann, die in den Punkten x_ν von vorgegebener Ordnung c_ν und sonst nirgends verschwindet (vgl. [6], p. 566). In der Funktionentheorie mehrerer Veränderlichen kann man in Analogie dazu folgenden Satz beweisen:

Satz 16. *Es sei $\mathfrak{G}=(G, \Phi)$ ein unverzweigtes Holomorphiegebiet; A sei eine rein $(n-1)$-dimensionale analytische Menge in G, die dort in die irreduziblen Bestandteile A_ν, $\nu=1, 2, 3, \ldots$, zerfallen möge. Ist dann $c_\nu>0$ eine beliebige Folge natürlicher Zahlen, so gibt es stets eine in G holomorphe, nicht identisch verschwindende Funktion $f(x)$, die in A_ν mindestens von der Ordnung c_ν verschwindet*[32]).

Dagegen kann man im allgemeinen nicht wie in der komplexen Analysis einer Veränderlichen erreichen, daß $f(x)$ auf A_ν genau von der Ordnung c_ν verschwindet und keine Nullstellen in $G-A$ hat[33]).

Eine meromorphe Funktion auf einer nicht-kompakten RIEMANNschen Fläche ist dort stets der Quotient zweier holomorpher Funktionen. Für die Holomorphiegebiete gilt entsprechend als Folge von Satz 16 die Aussage:

Satz 17. *Der Körper $K(G)$ der in einem zusammenhängenden, unverzweigten Holomorphiegebiet $\mathfrak{G}=(G, \Phi)$ meromorphen Funktionen ist der Quotientenkörper des Integritätsringes $R(G)$ der in G holomorphen Funktionen.*

Jedoch kann man im allgemeinen nicht wie in der klassischen Theorie jede meromorphe Funktion als Quotient zweier holomorpher Funktionen darstellen, die lokal teilerfremd sind[34]).

§ 4. Nicht-holomorph-konvexe Holomorphiegebiete

1. Über die Funktionentheorie in verzweigten RIEMANNschen Gebieten ist bisher wenig bekannt geworden. Es ist z.B. noch ein offenes Problem, ob jedes pseudokonvexe RIEMANNsche Gebiet ein Holomorphiegebiet oder sogar

[31]) Nach [19], Satz B, sind holomorph-konvexe, unverzweigte RIEMANNsche Gebiete über dem C sog. holomorph-vollständige Mannigfaltigkeiten (variétés de STEIN). Satz 15 ergibt sich daher aus den Resultaten von H. CARTAN [14]. Für schlichte Holomorphiegebiete wurde der Satz bereits 1937 von K. OKA [29] bewiesen.

[32]) In Verbindung mit [19], Satz B, ist Satz 16 eine direkte Folgerung aus [14], théorème A.

[33]) K. STEIN hat hierfür ein Beispiel gegeben in seiner Habilitationsschrift: Topologische Bedingungen für die Existenz analytischer Funktionen komplexer Veränderlichen zu vorgegebenen Nullstellenflächen. Math. Ann. **117**, 727—757 (1940/41).

[34]) Zwischen den Sätzen 15, 16, 17 bestehen einige Beziehungen. Vgl. dazu etwa [3].

holomorph-konvex ist. Wir werden jedoch in diesem Paragraphen ein verzweigtes Holomorphiegebiet konstruieren, das weder holomorph-konvex noch pseudokonvex ist. Wir ziehen dazu die in § 2.6 diskutierte analytisch-verzweigte Überlagerung (Y^n, Φ, C^n) heran. Den Raum Y^n denken wir uns wie in § 2.4 durch die Schar der Ebenen $E(s)$ überdeckt, wobei s in einer RIEMANNschen Zahlenkugel läuft.

Es sei nun K ein Kreis $\{|s|<d<1\}$ der s-Ebene. Offenbar ist dann die Menge $G^n = \bigcup_{s\in K} E(s) - y_0$ eine offene zusammenhängende Teilmenge von Y^n und $\mathfrak{G}_0^n = (G^n, \Phi)$ ein RIEMANNsches Gebiet über dem C^n, das höchstens $n-1$ Blätter hat. — Wir notieren einige Eigenschaften von $\mathfrak{G}_0^n$:

(1) *G^n ist analytisch isomorph dem kartesischen Produkt* $B^n = K\times\{C^k - O\}$, $k = n-1$.

Beweis. Die Abbildung $\pi: s = \frac{z_1}{w_1} = \frac{z_2}{w_2} = \cdots = \frac{z_k}{w_k}$, $v_1 = w_1, \ldots, v_k = w_k$ bildet G^n umkehrbar holomorph auf B^n ab.

Offenbar folgt aus (1), daß G^n eine komplexe Mannigfaltigkeit ist und somit nur uniformisierbare Punkte enthält. Wir zeigen weiter:

(2) *Es gibt einen Punkt* $\mathfrak{v}_0 = (v_0, v_1, \ldots, v_k) \in C^n$, *über dem genau ein Nichtverzweigungspunkt und kein weiterer Punkt von G^n liegt.*

Beweis. Wir wählen $\mathfrak{v}_0$ so, daß $\mathfrak{v}_0 \notin \Phi(E(\infty))$ und $P(s) = v_0 \prod_{\varkappa=1}^{k}(1+a_\varkappa s) - s\sum_{\varkappa=1}^{k} v_\varkappa \cdot \prod_{\substack{\nu=1\\ \nu\neq\varkappa}}^{k}(1+a_\nu s)$ mit den Werten $v_0, v_1, \ldots, v_k$ genau eine einfache Nullstelle $s_0 \in K$ und keine weiteren Nullstellen in $\overline{K}$ besitzt. Das ist möglich, da bei beliebigen $v_\varkappa$ auch die Koeffizienten von $P(s)$ beliebig sind. Zu s_0 gehört genau ein Punkt $x_0 \in E(s_0) - y_0 < G^n$, der über $\mathfrak{v}_0$ liegt. Es gibt über $\mathfrak{v}_0$ keine weiteren Punkte $x \in G^n$. Ferner gibt es aus Stetigkeitsgründen zu $\mathfrak{v}_0$ eine Umgebung $U(\mathfrak{v}_0)$, so daß über jedem Punkt $\mathfrak{v} \in U$ genau ein Punkt $x \in G^n$ liegt. G^n ist also in einer Umgebung von x_0 sicher unverzweigt, d.h. $\mathfrak{v}_0$ hat die in (2) behauptete Eigenschaft.

Wir fügen jetzt zu G^n alle erreichbaren Randpunkte r hinzu. Da G^n Teilgebiet von $Y^n - y_0$ ist und über $Y^n - y_0$ einen glatten Rand besitzt, läßt sich ∂G^n außerhalb y_0 durch die Häufungsmenge von G^n in $Y^n - y_0 - G^n$ repräsentieren. Die oben angegebene Abbildung π kann man sich in diese Menge fortgesetzt denken. Zu jedem Punkt $r \in \partial G^n \cap (Y^n - y_0)$ gibt es eine Umgebung $U(r) < Y^n - y_0$, die vermöge π umkehrbar holomorph auf eine Umgebung $U^*(r^*) < C^n$, $r^* = \pi(r) \in \partial K \times \{C^k - O\}$ abgebildet wird. ∂G^n hat also in der Nähe von r die gleiche analytische Struktur wie ∂B^n in der Nähe von r^*. Daraus ergibt sich:

(3) *Ist* $f^*(s, v_1, \ldots, v_k)$ *eine in B^n holomorphe Funktion, die in* $r^* = \pi(r)$, $r \in \partial G^n \cap (Y^n - y_0)$, *eine Singularität hat, so läßt sich auch* $f(x) = f^* \circ \pi(x)$ *nicht über* $r \in \partial G^n$ *hinaus fortsetzen.*

Wir beweisen ferner:

(4) *Über* $O \in C^n$ *liegt genau ein Randpunkt* $r_0 \in \partial G^n$.

Beweis. Es genügt zu zeigen, daß das Urbild $\Phi^{-1}(H)$ jeder Hyperkugel $H(O)$ eine nicht leere zusammenhängende Menge in G^n ist. Das aber ist klar, da $\pi \circ \Phi^{-1}(H) = \{(s(v_1 + \cdots + v_k),\ v_1(1+a_1 s), \ldots, v_k(1+a_k s)) \in H\} \cap B^n$ immer eine zusammenhängende Menge in B^n um $K \times O$ ist.

Dem Kontinuum der Randpunkte $K \times O$ von B^n entspricht also genau der Randpunkt r_0 von $\mathfrak{G}_0^n$, während die Punkte von $\partial K \times (C^k - O)$ und $\partial G^n - \breve{\Phi}^{-1}(O)$ eineindeutig einander zugeordnet sind. Man kann π^{-1} zu einer stetigen Abbildung $\bar{\pi}^{-1}$ von $\bar{B}^n$ auf $\breve{G}^n$ fortsetzen.

2. Wir zeigen nun:

(5) *$\mathfrak{G}_0^n$ ist ein Holomorphiegebiet.*

Beweis. $K \times C^k$ ist als Polyzylinder ein Holomorphiegebiet ([2], p. 36ff.). Es gibt also eine holomorphe Funktion $f^*(\mathfrak{v})$, die $K \times C^k$ zum Existenzgebiet hat. Nach (3) ist $f(x) = f^* \circ \pi(x)$ eine in G^n holomorphe Funktion, die in allen Randpunkten $r \in \partial G^n - r_0$ singulär ist. r_0 selbst ist aber Häufungspunkt dieser Randpunkte; denn $\bar{\pi}^{-1}$ bildet $\bar{B}^n$ stetig auf $\breve{G}^n$ ab, wobei das Urbild jeder Umgebung von r_0 eine volle Umgebung $U(\bar{K} \times O) \subset \bar{B}^n$ umfaßt. U enthält stets Punkte von $\partial K \times (C^k - O)$. Es folgt also, daß $f(x)$ auch in r_0 wesentlich singulär wird.

Es sei nun $\mathfrak{G}^* = (G^*, \Phi^*)$ ein Existenzgebiet von $f(x)$. $\mathfrak{G}_0^n$ ist in $\mathfrak{G}^*$ enthalten, $f(x)$ ist in G^* hinein fortsetzbar. Die spurpunkttreue Abbildung $\tau: G^n \to G^*$ ist zu einer stetigen Abbildung $\breve{\tau}: \breve{G}^n \to \breve{G}^*$ fortsetzbar. $\breve{\tau}$ bildet keinen Punkt von ∂G^n in das Innere von G^* ab, da $f(x)$ überall in ∂G^n singulär ist. Folglich ist $\mathfrak{A} = (G^n, \tau, G^*)$ eine unbegrenzte Überlagerung von G^*. Da τ außerdem spurpunkttreu abbildet und nach (2) über einem gewissen $\mathfrak{v}_0 \in C^n$ genau ein Nichtverzweigungspunkt von G^n liegt, muß $\mathfrak{A}$ einblättrig sein: τ bildet mithin G^n umkehrbar eindeutig auf G^* ab. $\mathfrak{G}_0^n$ und $\mathfrak{G}^*$ sind identisch, d.h. $\mathfrak{G}_0^n$ ist selbst das Existenzgebiet von $f(x)$ und damit ein Holomorphiegebiet.

Weiter gilt:

(6) *$\mathfrak{G}_0^n$ ist in allen Randpunkten $r \neq r_0$ pseudokonvex. In r_0 ist dagegen $\mathfrak{G}_0^n$ nicht pseudokonvex.*

Beweis. Zunächst ist $\mathfrak{G}_0^n$ pseudokonvex in jedem Randpunkt $r \neq r_0$, da in der Nähe eines solchen Punktes ∂G^n die gleiche Struktur hat wie ∂B^n in der Nähe von $\pi(r)$. ∂B^n ist aber dort der Rand eines Polyzylinders.

Demgegenüber ist $C^k - O$, $k \geq 2$, in O nicht pseudokonvex. Es gibt vielmehr zu jedem $\varepsilon > 0$ eine in $C^k - O$ ausgezeichnete Schar ${}^*\sigma_\varepsilon(w, t)$, die O enthält. Für festes $s_0 \in K$ ist dann $(s_0, {}^*\sigma_\varepsilon(w, t))$ eine ausgezeichnete Schar in $B^n = K \times (C^k - O)$, deren Durchschnitt mit $K \times O$ nicht leer ist. Offenbar enthält die ausgezeichnete Schar $\sigma_\delta(w, t) = \bar{\pi}^{-1}(s_0, {}^*\sigma_\varepsilon(w, t))$ in $\mathfrak{G}_0^n$ den Punkt r_0. Da $\delta > 0$ beliebig klein wird, wenn ε gegen 0 geht, haben wir gezeigt, daß $\mathfrak{G}_0^n$ in r_0 nicht pseudokonvex ist, q. e. d.

Die Holomorphiehülle von B^n ist $K \times C^k \neq B^n$. Daher ist B^n sicher nicht holomorph-konvex (vgl. [2], p. 73). Da ferner G^n und B^n analytisch isomorph sind und die Holomorphiekonvexität gegenüber eineindeutigen holomorphen Abbildungen invariant ist, folgt:

(7) *G^n ist nicht holomorph-konvex.*

Zusammenfassend ist also bewiesen:

Satz 18. *Es gibt über dem C^n, $n \geq 3$, endlich-blättrige verzweigte Holomorphiegebiete, die aus lauter uniformisierbaren Punkten bestehen und weder holomorph-konvex noch pseudokonvex sind.*

3. Es ist bisweilen vermutet worden, daß die Holomorphiegebiete gegenüber holomorphen Transformationen invariant sind. Damit ist folgendes gemeint:

Sind $\mathfrak{G}_1 = (G_1, \Phi_1)$ und $\mathfrak{G}_2 = (G_2, \Phi_2)$ zwei RIEMANNsche *Gebiete über dem C^n und ist φ eine umkehrbar holomorphe Abbildung von G_1 auf G_2, so ist $\mathfrak{G}_1$ dann und nur dann ein Holomorphiegebiet, wenn $\mathfrak{G}_2$ ein Holomorphiegebiet ist.*

Für unverzweigte RIEMANNsche Gebiete ist die Richtigkeit dieser Vermutung bestätigt worden; denn in diesem Falle werden die Holomorphiegebiete durch die Holomorphiekonvexität charakterisiert (vgl. Satz 11). Die Holomorphiekonvexität hat aber auf Grund ihrer Definition die angegebene Invarianzeigenschaft.

Unser oben konstruiertes RIEMANNsches Gebiet $\mathfrak{G}_0^n = (G^n, \Phi)$ lehrt nun, daß für verzweigte RIEMANNsche Gebiete die Vermutung falsch ist. Wie wir gezeigt haben, ist G^n vermöge π umkehrbar holomorph auf B^n abgebildet. Bezeichnet ι die Injektion $B^n \to C^n$, so ist das RIEMANNsche Gebiet $\mathfrak{B}^n = (B^n, \iota)$ dem Holomorphiegebiet $\mathfrak{G}_0^n$ analytisch äquivalent; indessen ist jedoch $\mathfrak{B}^n$, $n \geq 3$, kein Holomorphiegebiet.

Es ist klar, daß funktionentheoretische Aussagen für $\mathfrak{G}_0^n$ und $\mathfrak{B}^n$ simultan gelten, wenn sie sich nur auf die komplexen Räume G^n bzw. B^n beziehen. Wir brauchen deshalb solche Aussagen nur für B^n nachzuweisen.

Nach CARTAN [11] gibt es im $C^2 - O$ eine Cousin-I-Verteilung $\{f_i^*\}$, die dort keine Lösung besitzt. Setzen wir $f_i(s, v_1, v_2) = f_i^*(v_1, v_2)$, so erhalten wir in B^3 eine Cousin-I-Verteilung, die in B^3 nicht lösbar ist. Damit ist bewiesen:

Satz 19. *Es gibt verzweigte Holomorphiegebiete über dem C^n, $n \geq 3$, in denen nicht jede Cousin-I-Verteilung eine Lösung besitzt.*

Die Aussagen der Sätze 16, 17 behalten im Gegensatz zur Aussage von Satz 15 in den Holomorphiegebieten $\mathfrak{G}_0^n = (G^n, \Phi)$ ihre Gültigkeit. Wie vorhin brauchen wir das nur für den Raum B^n zu zeigen. Ist etwa A eine rein $(n-1)$-dimensionale analytische Menge in $B^n = K \times (C^k - O)$, $n \geq 3$, die dort in die irreduziblen Komponenten A_ν, $\nu = 1, 2, \ldots$, zerfällt, so läßt sich A_ν nach einem Fortsetzungssatz von THULLEN [40] zu einer in $K \times C^k$ analytischen Menge A'_ν fortsetzen. Da $K \times C^k$ ein Holomorphiegebiet im C^n ist, folgt aus Satz 16, daß man zu jeder Folge natürlicher Zahlen $c_\nu > 0$ eine in $K \times C^k$ holomorphe Funktion $f \not\equiv 0$ finden kann, die auf A'_ν mindestens von der Vielfachheit c_ν

verschwindet. Da f erst recht in B^n holomorph ist, ergibt sich, daß f auch in B^n die Aussage von Satz 16 befriedigt. Satz 16 gilt also auch in $\mathfrak{G}_0^n$ [35]).

Unter Benutzung eines Fortsetzungssatzes für meromorphe Funktionen kann man in analoger Weise zeigen, daß ebenfalls Satz 17 seine Gültigkeit in $\mathfrak{G}_0^n$ behält. Es dürfte interessieren, ob Satz 16 und Satz 17 uneingeschränkt für beliebige Holomorphiegebiete gelten.

In dieser Arbeit wurden für die Dimensionen $n \geq 3$ verzweigte Holomorphiegebiete angegeben, die nicht holomorph-konvex sind. Im Falle $n = 2$ versagt das angegebene Konstruktionsverfahren. Es bleibt offen, ob es auch nicht-holomorph-konvexe Holomorphiegebiete über dem C^2 gibt.

Literatur

[1] BEHNKE, H.: Die analytischen Gebilde von holomorphen Funktionen mehrerer Veränderlichen. Arch. d. Math. **6**, 353—368 (1955). — [2] BEHNKE, H., u. P. THULLEN: Theorie der Funktionen mehrerer komplexer Veränderlichen. Erg. d. Math. **3**. Berlin 1934. — [3] BEHNKE, H., u. K. STEIN: Analytische Funktionen mehrerer Veränderlichen zu vorgegebenen Null- und Polstellenflächen. Jber. DMV. **47**, 177—192 (1937). — [4] BEHNKE, H., u. K. STEIN: Konvergente Folgen nichtschlichter Regularitätsbereiche. Ann. Mat. pura appl. **28**, 317—326 (1949). — [5] BEHNKE, H., u. K. STEIN: Modifikation komplexer Mannigfaltigkeiten und Riemannscher Gebiete. Math. Ann. **124**, 1—16 (1951). — [6] BEHNKE, H., u. F. SOMMER: Theorie der analytischen Funktionen einer komplexen Veränderlichen, Lehrbuch. Berlin 1955. — [7] BLACK, C. W. M.: The parametric representation of the neighborhood of a singular point of an analytic surface. Proc. Amer. Acad. **37** (1901). — [8] CARATHÉODORY, C.: Über die analytischen Abbildungen von mehrdimensionalen Räumen. Verh. Int. Math. Congr. Zürich 1932, vol. I, 93—101. — [9] CARTAN, H.: Sur les domaines d'existence des fonctions de plusieurs variables complexes. Bull. Soc. Math. France **59**, 46—69 (1931). — [10] CARTAN, H., u. P. THULLEN: Zur Theorie der Singularitäten der Funktionen mehrerer komplexer Veränderlichen. Regularitäts- und Konvergenzbereiche. Math. Ann. **106**, 617—647 (1932). — [11] CARTAN, H.: Les problèmes de Poincaré et de Cousin pour les fonctions de plusieurs variables complexes. C. R. Acad. Sci., Paris **199** (1934). — [12] CARTAN, H.: Problèmes globaux dans la théorie des fonctions analytiques de plusieurs complexes. Proc. Int. Congr. Math. 1950, vol. I, 152—164. — [13] CARTAN, H.: Séminaire E.N.S., Paris 1951/52 (hektographiert). — [14] CARTAN, H.: Variétés analytiques complexes et cohomologie; Coll. sur les Fonct. de Plus. Variables, Bruxelles 1953, 41—55. — [15] CARTAN, H.: Séminaire E.N.S., Paris 1953/54 (hektographiert). — [16] COUSIN, P.: Sur les fonctions de n variables complexes. Acta Math. **19**, 1—62 (1895). — [17] EHRESMANN, CH.: Les prolongements d'une variété différentiable. AH. d. IV. Congr. dell'Un. Math. Ital., Taormina 1951. — [18] EHRESMANN, CH.: Struktures locales; Conférence à Rome, März 1952 (hektographiert). — [19] GRAUERT, H.: Charakterisierung der holomorph vollständigen komplexen Räume. Math. Ann. **129**, 233—259 (1955). — [20] GRAUERT, H.: Charakterisierung der Holomorphiegebiete durch die Kählersche Metrik. Math. Ann. **131** (1956). — [21] GRAUERT, H., u. R. REMMERT: Zur Theorie der Modifikationen I. Stetige und eigentliche Modifikationen komplexer Räume. Math. Ann. **129**, 274—296 (1955). — [22] GRAUERT, H., u. R. REMMERT: Konvexität in der komplexen Analysis. Nicht holomorph-

[35]) Unter Benutzung des veralligemeinerten THULLENschen Singularitätensatzes (vgl. [35]) kann man sogar zeigen:

Jede im Holomorphiegebiet $\mathfrak{G}_0$ rein q-dimensionale analytische Menge ist das genaue Nullstellengebilde von höchstens $n+1$ in $\mathfrak{G}_0$ holomorphen Funktionen, wenn $q > 1$. Dagegen gibt es eindimensionale analytische Mengen in $\mathfrak{G}_0$, die sich nicht als gemeinsame Nullstellenmenge von in $\mathfrak{G}_0$ holomorphen Funktionen — auch nicht von unendlich vielen — darstellen lassen.

konvexe Holomorphiegebiete und Anwendungen auf die Abbildungstheorie. Com. Math. Helv. 1956/57. — [23] GRAUERT, H., u. R. REMMERT: Analytisch-verzweigte Überlagerungen und komplexe Räume (in Vorbereitung). — [24] HIRZEBRUCH, F.: Über vierdimensionale Riemannsche Flächen mehrdeutiger analytischer Funktionen von zwei komplexen Veränderlichen. Math. Ann. **126**, 1—22 (1953). — [25] HODGE, W. V. D., u. D. PEDOE: Methods of Algebraic Geometry III. Cambridge 1954. — [26] HOPF, H.: Zur Topologie der komplexen Mannigfaltigkeiten. Studies and Essays presented to R. COURANT, New York 167—185 (1948). — [27] JUNG, H. W. E.: Darstellung der Funktionen eines algebraischen Körpers zweier unabhängiger Veränderlichen in der Umgebung einer Stelle $x = a$, $y = b$. J. reine u. angew. Math. **133**, 289—314 (1908). — [28] MENGER, K.: Diemensionstheorie. Leipzig 1928. — [29] OKA, K.: Sur les fonctions analytiques de plusieurs variables. II—Domaines d'holomorphie. J. of Sci. Hirosima Univ. Ser. A **7**, Nr. 2, 115—130 (1937). — [30] OKA, K.: Sur les fonctions analytiques de plusieurs variables. IX—Domaines finis sans point critique intérieur. Jap. J. Math. **23**, 97—155 (1953). — [31] OSGOOD, W. F.: Lehrbuch der Funktionentheorie, II, 1. Berlin u. Leipzig 1929. — [32] REMMERT, R.: Über stetige und eigentliche Modifikationen komplexer Räume. Coll. de Topologie, Strasbourg 1954. — [33] REMMERT, R.: Projektionen analytischer Mengen. Math. Ann. **130**, 410—441 (1956). — [34] REMMERT, R.: Holomorphe und meromorphe Abbildungen komplexer Räume. Math. Ann. **1957**. — [35] REMMERT, R., u. K. STEIN: Über die wesentlichen Singularitäten analytischer Mengen. Math. Ann. **126**, 263—306 (1953). — [36] STEIN, K.: Analytische Abbildungen allgemeiner analytischer Räume. Coll. de Topologie, Strasbourg 1954. — [37] STEIN, K.: Analytische Zerlegungen komplexer Räume II (in Vorbereitung). — [38] TEICHMÜLLER, O.: Veränderliche Riemannsche Flächen. Dtsch. Math. **1944**, 344—359. — [39] THULLEN, P.: Zur Theorie der Singularitäten der Funktionen zweier komplexer Veränderlichen. Die Regularitätshüllen. Math. Ann. **106**, 64—76 (1932). — [40] THULLEN, P.: Über die wesentlichen Singularitäten analytischer Funktionen und Flächen im Raume von n komplexen Veränderlichen. Math. Ann. **111**, 137—157 (1935). — [41] WAERDEN, B. L. VAN DER: Die Bedeutung des Bewertungsbegriffs für die algebraische Geometrie. Jber. Math. Ver. **52**, 161—172 (1942). — [42] WEIERSTRASS, K.: Allgemeine Untersuchungen über $2n$-fach periodische Funktionen von n Veränderlichen. Wk. III, 53—114. — [43] ZARISKI, O.: Reduction of the singularities of algebraic three dimensional varieties. Ann. of Math. **45**, 472—542 (1944).

Münster, Mathematisches Institut der Universität

(Eingegangen am 24. März 1956)

GRUNSKY, H.
Math. Zeitschr. Bd. 67, S. 129—132 (1957)

Über konforme Abbildungen, die gewisse Gebietsfunktionen in elementare Funktionen transformieren. I

Herrn GEORG FABER zum 80. Geburtstage gewidmet

Von

HELMUT GRUNSKY

Herr J. L. WALSH hat kürzlich zwei Sätze über konforme Abbildung endlich vielfach zusammenhängender Gebiete bewiesen[1]). Der erste besagt, daß stets ein Bildgebiet gewonnen werden kann, das berandet wird von zwei vollständigen Niveaulinien $|S(w)|=1$ und $|S(w)|=k>1$ einer Funktion

$$S(w)=A\,\frac{\prod_{\mu=1}^{m}(w-a_\mu)^{M_\mu}}{\prod_{\nu=1}^{n}(w-b_\nu)^{N_\nu}},\qquad \sum M_\mu=\sum N_\nu,$$

wo für die M_μ und N_ν beliebige reelle Werte zugelassen sind. Dabei können die beiden Teilsysteme der Randkomponenten, auf denen $|S|=1$ bzw. $|S|=k$ gilt, willkürlich gewählt werden. Setzt man ohne Beschränkung der Allgemeinheit $k=e$ (indem man von der verbleibenden Freiheit in der Wahl der Konstanten A, M_μ, N_ν Gebrauch macht), so heißt das: Das harmonische Maß $h_c(z)$ eines beliebig herausgegriffenen Teilsystems C der Randkomponenten läßt sich durch konforme Abbildung in eine elementare Funktion

$$\tilde{h}_c(w)=\sum_\mu M_\mu \log|w-a_\mu|-\sum_\nu N_\nu \log|w-b_\nu|+\text{const}$$

transformieren. Der zweite Satz läßt sich ähnlich interpretieren als Transformation von $\sum_{\varkappa=1}^{k}\alpha_\varkappa g(z,z_\varkappa)$ in eine elementare Funktion, wo $g(z,z_\varkappa)$ die GREENsche Funktion des gegebenen Gebietes mit Singularität $z_\varkappa$ ist und die $\alpha_\varkappa>0$ sind[2]).

[1]) WALSH, J. L.: C. r. Acad. Sci., Paris **239**, 1572—1574, 1756—1758 (1954).

[2]) Die Untersuchungen von WALSH knüpfen inhaltlich und methodisch an Arbeiten von DE LA VALLÉE POUSSIN und JULIA aus den Jahren 1930—1932 an, bei denen es sich jedoch vorwiegend um die komplizierteren Probleme handelte, die entstehen, wenn S eine rationale, speziell eine ganz rationale Funktion ist, während der Rand des Bildes nicht notwendig aus vollständigen Niveaulinien besteht; vgl. G. JULIA, Leçons sur la représentation conforme des aires multiplement connexes, Paris 1934. Nur eine Note von DE LA VALLÉE POUSSIN bezieht sich auf einen Spezialfall des Problems von WALSH: C. r. Acad. Sci., Paris **192**, 128—131 (1931).

Zusatz bei der Korrektur: In einem zweiten Teil wird die hier verwendete Methode so weitergebildet, daß sie auch den Beweis der andern Ergebnisse von DE LA VALLÉE POUSSIN und JULIA ergibt.

Unter Benutzung eines grundlegenden Satzes der Uniformisierungstheorie gebe ich in dieser Note einen sehr kurzen Beweis eines etwas allgemeineren Satzes.

$\mathfrak{Z}$ sei ein n-fach zusammenhängendes Gebiet der z-Ebene, sein Rand $Z = \bigcup_{\nu=1}^{n} Z_\nu$. Keine der Komponenten Z_ν sei in einen Punkt ausgeartet, und sie mögen, ohne Beschränkung der Allgemeinheit, als analytische JORDAN-Kurven vorausgesetzt werden. Endlich sei $\infty \in \mathfrak{Z}$ angenommen. $z_\varkappa (\varkappa = 1, \ldots, k)$ sei ein gegebenes System von Punkten aus $\mathfrak{Z}$, die wir als $\neq \infty$ annehmen wollen, $\alpha_\varkappa (\varkappa = 1, \ldots, k)$, $c_\nu (\nu = 1, \ldots, n)$ zwei Systeme von reellen Zahlen. $\mathfrak{g}(z)$ sei eine reellwertige Funktion für $z \in \mathfrak{Z}$ mit folgenden Eigenschaften: 1. $\mathfrak{g}(z) + \sum_{\varkappa=1}^{k} \alpha_\varkappa \log |z - z_\varkappa|$ ist harmonisch in $\mathfrak{Z}$. 2. $\mathfrak{g}(z)$ hat auf Z_ν die konstanten Randwerte c_ν. 3. Die Ableitung nach der äußeren Randnormalen, $\frac{\partial \mathfrak{g}}{\partial n}$, wechselt auf keinem Z_ν das Vorzeichen (es ist also für verschiedene Z_ν verschiedenes Vorzeichen zugelassen, und auch $\frac{\partial \mathfrak{g}}{\partial n} = 0$ ist nicht ausgeschlossen, kann aber natürlich nur in isolierten Punkten vorkommen). Bezeichnet $g(z, \zeta)$ die GREENsche Funktion von $\mathfrak{Z}$, $h_\nu(z)$ das harmonische Maß von Z_ν, so ist

$$\mathfrak{g}(z) = \sum_{\varkappa=1}^{k} \alpha_\varkappa g(z, z_\varkappa) + \sum_{\nu=1}^{n} c_\nu h_\nu(z).$$

Die Bedingung 3. bedeutet eine Einschränkung für die Wahl der Konstanten $z_\varkappa, \alpha_\varkappa, c_\nu$. Sie ist insbesondere erfüllt, wenn alle $\alpha_\varkappa = 0$ und für die c_ν nur zwei verschiedene Werte zugelassen sind; ebenso, wenn alle $c_\nu = 0$ und die $\alpha_\varkappa$ alle von einerlei Vorzeichen sind. Diese beiden Spezialisierungen führen auf die eingangs genannten Sätze von Herrn WALSH.

Wir beweisen nun den

Satz 1. *Es gibt genau eine konforme Abbildung $w = F(z)$ von $\mathfrak{Z}$, die bei ∞ durch*

$$F(z) = z + O\left(\frac{1}{z}\right) \tag{1}$$

normiert ist, derart, daß sich die Funktion $\mathfrak{g}(z)$ transformiert in $\tilde{\mathfrak{g}}(w) = \log |S(w)|$ mit

$$\begin{cases} S(w) = A \prod_{\varkappa=1}^{k} (w - w_\varkappa)^{-\alpha_\varkappa} \prod_{\nu=1}^{n} (w - b_\nu)^{-\beta_\nu} \\ A = e^{\mathfrak{g}(\infty)},\ w_\varkappa = F(z_\varkappa),\ \beta_\nu = \frac{1}{2\pi} \int_{Z_\nu} \frac{\partial \mathfrak{g}}{\partial n}\, ds, \end{cases} \tag{2}$$

wobei in jeder Komponenten der Komplementärmenge des Bildgebietes genau ein Punkt b_ν liegt[3]).

Beweis. Es sei $\mathfrak{Z}_\nu$ das Innere von Z_ν, a_ν ein beliebiger fester Punkt aus $\mathfrak{Z}_\nu$, $g_\nu(z, \zeta)$ die GREENsche Funktion von $\mathfrak{Z}_\nu$. Wir erklären auf Z_ν (mit der

[3]) Man beachte die Gl. (6), aus der die Regularität von $S(w)$ in ∞ und ihre Invarianz gegenüber einer Streckung folgt.

Bogenlänge s) zwei Funktionen, indem wir zunächst willkürlich einen Anfangspunkt $z_{\nu 0} \in Z_\nu$ wählen und setzen:

$$\varphi(z) = \frac{1}{2\pi} \int_{z_{\nu 0}}^{z} \frac{\partial \mathfrak{g}}{\partial n}\, ds, \qquad \varphi_\nu(z) = \frac{\beta_\nu}{2\pi} \int_{z_{\nu 0}}^{z} \frac{\partial g_\nu(z, a_\nu)}{\partial n}\, ds, \qquad z \in Z_\nu\ ^{4}).$$

Beide Funktionen ändern sich bei einem Umlauf um Z_ν monoton im strengen Sinne von 0 bis β_ν. Daher wird durch die Gleichung

$$\varphi(z) = \varphi_\nu(z^{(\nu)}) \tag{3}$$

eine umkehrbar eindeutige Abbildung von Z_ν auf sich selbst erklärt; sie ist in beiden Richtungen regulär analytisch, abgesehen von den endlich vielen Nullstellen von $\frac{\partial \mathfrak{g}}{\partial n}$. Indem wir nun einen Punkt $z^{(\nu)}$ von Z_ν als Randpunkt von $\mathfrak{Z}_\nu$ identifizieren mit demjenigen Randpunkt z von $\mathfrak{Z}$, der ihm vermöge (3) entspricht, machen wir die z-Ebene zu einer RIEMANNschen Mannigfaltigkeit vom topologischen Typus der Vollkugel. Sie läßt sich umkehrbar eindeutig und konform auf eine w-Kugel abbilden, und wir normieren die vermittelnde Funktion $w = F(z)$ gemäß (1). Die Bilder der $z_\varkappa$ seien mit $w_\varkappa$, die der a_ν mit b_ν bezeichnet. Dann ist leicht zu sehen, daß bei dieser Abbildung die Transformierte von $\mathfrak{g}(z)$ in der Form (2) erscheint.

Bezeichnet man nämlich die durch Verpflanzung einer Funktion aus der z- in die w-Ebene entstehende Funktion durch das zusätzliche Zeichen $\sim$, so ist $\tilde{g}_\nu(w, b_\nu)$ die GREENsche Funktion von $\mathfrak{W}_\nu$, dem Inneren des Bildes W_ν von Z_ν, und wegen der Invarianz von $\frac{\partial g_\nu}{\partial n}\, ds$ und $\frac{\partial \mathfrak{g}}{\partial n}\, ds$ auf dem Rande gilt in einem beliebigen Punkte $w = F(z) = w^{(\nu)} = F(z^{(\nu)}) \in W_\nu$:

$$\tilde{\varphi}(w) = \frac{1}{2\pi} \int_{w_{\nu 0}}^{w} \frac{\partial \tilde{\mathfrak{g}}}{\partial n}\, ds = \tilde{\varphi}(w^{(\nu)}) = \frac{\beta_\nu}{2\pi} \int_{w_{\nu 0}}^{w^{(\nu)}} \frac{\partial \tilde{g}_\nu}{\partial n}\, ds, \qquad w_{\nu 0} = F(z_{\nu 0});$$

daher gilt auf W_ν:

$$\frac{\partial \tilde{\mathfrak{g}}}{\partial n} = \beta_\nu \frac{\partial \tilde{g}_\nu}{\partial n}; \tag{4}$$

außerdem ist dort

$$\frac{\partial \tilde{\mathfrak{g}}}{\partial s} = \beta_\nu \frac{\partial \tilde{g}_\nu}{\partial s} = 0. \tag{5}$$

Nun ist $\frac{\partial \tilde{\mathfrak{g}}(w)}{\partial w}$ [5]), abgesehen von den $w_\varkappa$, eine analytische Funktion von w,

[4]) $\frac{\partial}{\partial n}$ bedeutet immer die Ableitung nach der *in bezug auf* $\mathfrak{Z}$ bzw. sein Bild $\mathfrak{W}$ äußeren Randnormalen.

[5]) $\frac{\partial}{\partial w}$ ist erklärt durch $\frac{\partial}{\partial w} = \frac{1}{2}\left(\frac{\partial}{\partial u} - i\,\frac{\partial}{\partial v}\right)$ $(w = u + iv)$. Falls die Funktion $p(w)$, auf die der Operator angewandt wird, reell ist, so ist $\frac{\partial p}{\partial w} = \frac{1}{2}\,\overline{\text{grad}\, p}$, und bei harmonischem p ist das eine analytische Funktion. In diesem Falle kann auch geschrieben werden $p = \frac{1}{2}(P + \bar{P})$, wo P eine analytische Funktion ist, und es ist $\frac{\partial p}{\partial w} = \frac{1}{2} P'$. Insbesondere ist also, wenn p in ∞ regulär harmonisch ist dort $\frac{\partial p}{\partial w} = O(w^{-2})$.

ebenso $\beta_\nu \frac{\partial \tilde{g}_\nu(w, b_\nu)}{\partial w}$, abgesehen von b_ν, und die beiden Gln. (4) und (5) besagen, daß diese beiden Funktionen auf W_ν übereinstimmende Randwerte haben und sonach analytische Fortsetzung voneinander sind. Die so definierte analytische Funktion $R(w)$ ist also in der ganzen w-Ebene regulär mit Ausnahme der $w_\varkappa$ und der b_ν. In diesen Punkten hat sie Pole mit den Hauptgliedern $\frac{1}{2}\frac{-\alpha_\varkappa}{w-w_\varkappa}$ bzw. $\frac{1}{2}\frac{-\beta_\nu}{w-b_\nu}$. Nun ist aber nach dem GAUSSschen Integralsatz

$$\sum_{\varkappa=1}^{k} \alpha_\varkappa + \sum_{\nu=1}^{n} \beta_\nu = 0. \tag{6}$$

Daher ist

$$R(w) = \frac{1}{2} \sum \frac{-\alpha_\varkappa}{w-w_\varkappa} + \frac{1}{2} \sum \frac{-\beta_\nu}{w-b_\nu}.$$

Denn die rechte Seite verschwindet in $w=\infty$ in zweiter Ordnung, wie das für $\frac{\partial \tilde{\mathfrak{g}}}{\partial w}$ der Fall sein muß. Integration ergibt

$$\tilde{\mathfrak{g}}(w) = 2\,\mathfrak{Re} \int R(w)\,dw = \mathfrak{Re}\left\{\sum \log(w-w_\varkappa)^{-\alpha_\varkappa} + \sum \log(w-b_\nu)^{-\beta_\nu} + c\right\},$$

womit die Existenzaussage des Satzes bewiesen ist.

Die Einzigkeitsaussage ergibt sich leicht nach folgender Überlegung: Angenommen, es gebe zwei Abbildungen, die das Verlangte leisten. Als Bildebene der zweiten wählen wir die z-Ebene. Da nun die Funktionen $\mathfrak{g}(z)$ und $\beta_\nu g_\nu(z, a_\nu)$ denselben analytischen Ausdruck haben, so ist $\frac{\partial \mathfrak{g}}{\partial n} = \beta_\nu \frac{\partial \mathfrak{g}_\nu}{\partial n}$, also $\varphi(z) \equiv \varphi_\nu(z)$, und die Gl. (3) identifiziert Punkte, die ohnehin zusammenfallen. Daher ist die oben eingeführte RIEMANNsche Mannigfaltigkeit die volle z-Ebene, $w = F(z)$ eine umkehrbar eindeutige Abbildung derselben und mit Rücksicht auf ihre Normierung die Identität.

Mainz, Mathematisches Institut der Universität

(Eingegangen am 31. Mai 1956)

9254-67 10 B

Math. Zeitschr. Bd. 67, S. 133—138 (1957)

Zur konformen Abbildung schlichter Gebiete

Herrn Geheimrat FABER zum 80. Geburtstag gewidmet

Von

J. HEINHOLD

Es soll im folgenden der Begriff der Schmiegungsfunktion für konforme Abbildungen, wie er von KOEBE [1] beim Beweis des RIEMANNschen Abbildungssatzes für einfach zusammenhängende Gebiete benutzt wurde, auf mehrfach zusammenhängende Gebiete verallgemeinert werden. Bei KOEBE wird das nach Voraussetzung im Innern des Einheitskreises gelegene einfach zusammenhängende schlichte Gebiet $\mathfrak{G}$ mit dem Nullpunkt als innerem Punkt durch eine Folge analytischer Funktionen $\varphi_\nu(z_\nu) = z_{\nu+1}$ mit der Normierung $\varphi_\nu(0) = 0$, $\varphi_\nu'(0) > 0$ auf eine Folge von Gebieten $\mathfrak{G}_{\nu+1}$, $\mathfrak{G}_0 = \mathfrak{G}$, $z_0 = z$ mit den Rändern $\mathfrak{R}_{\nu+1}$ umkehrbar eindeutig und konform so abgebildet, daß der Minimalabstand der Randpunkte von $\mathfrak{R}_{\nu+1}$ vom Nullpunkt mit $\nu \to \infty$ monoton gegen 1 geht. Es konvergiert dann

$$\Phi_\nu(z) = \varphi_\nu\big(\varphi_{\nu-1}\big(\ldots(\varphi_0(z))\ldots\big)\big)$$

für $\nu \to \infty$ gleichmäßig in $\mathfrak{G}$ gegen die gesuchte Abbildungsfunktion. Die Funktionen φ_ν nennt KOEBE Schmiegungsfunktionen. Er verwendet hierfür die bekannten schon von CARATHÉODORY benutzten Quadratwurzelabbildungen. Für die Praxis brauchbare, den Besonderheiten der vorliegenden Gebiete anzupassende Schmiegungsfunktionen für die Abbildung ein- bzw. zweifach zusammenhängender Gebiete wurden von RINGLEB [2], HEINHOLD [3], [4] und ALBRECHT [4], [5] bzw. [6] angegeben.

1.

Wie seit den grundlegenden Arbeiten KOEBEs [7] bekannt ist, läßt sich jedes n-fach zusammenhängende schlichte Gebiet $\mathfrak{G}$ durch eine in $\mathfrak{G}$ analytische schlichte Funktion auf gewisse Typen von Normalgebieten $\mathfrak{N}$ abbilden, etwa auf Geradenschlitzgebiete mit parallelen oder radial vom Nullpunkt aus verlaufenden Schlitzen oder auf Kreisbogenschlitzgebiete, z.B. auf die längs konzentrischer Kreisbogen um den Nullpunkt geschlitzte Ebene, bzw. Einheitskreis oder konzentrischen Kreisring um den Nullpunkt. Die Abbildungsfunktionen sind unter gewissen Normierungsbedingungen eindeutig festgelegt und können als Lösungen gewisser Extremalprobleme innerhalb der Familie der in dem abzubildenden Gebiet regulären schlichten Funktionen betrachtet werden. Man kann (s. z.B. [8]) in jedem dieser Fälle ein stetiges Funktional $F(f)$ dieser in $\mathfrak{G}$ regulären schlichten Funktionen $f(z)$ angeben, dessen Betrag für die gesuchte Abbildungsfunktion auf das Normalgebiet ein Extremum annimmt. Mitunter stehen mehrere derartige Funktionale

zur Verfügung. Wir werden im folgenden nur beschränkte Funktionale voraussetzen, was für die Konstruktion praktisch brauchbarer Schmiegungsverfahren von Bedeutung sein kann (s. Fußnote 3).

Es liege ein n-fach zusammenhängendes, von n Randkontinuen mit jeweils mehr als einem Randpunkt begrenztes, gemäß gewissen Bedingungen (B) normiertes Gebiet $\mathfrak{G}$ vor. Hierunter sei insbesondere die Bedingung, daß $\mathfrak{G}$ in einem vorgegebenem endlichen Gebiet (z.B. im Falle $n=1$ bzw. $n \geq 2$ im Innern des Einheitskreises bzw. eines konzentrischen Kreisringes um den Nullpunkt) gelegen sei. Eine in $\mathfrak{G}$ reguläre, schlichte Abbildungsfunktion, welche die Bedingungen (B) unverändert läßt, nennen wir eine „nach (B) normierte Funktion". Von den Bedingungen (B) kann weiter vorausgesetzt werden, daß eine hiernach normierte Funktion keine Konstante sein und eine konvergente Folge (B)-normierter Funktionen nur wieder gegen eine (B)-normierte Funktion konvergieren kann. Die nach (B) normierten Funktionen bilden dann eine normale, kompakte Familie $\mathfrak{F}$.

Def. 1. $F(f)$ sei ein für die Funktionen $f(z)$ von $\mathfrak{F}$ definiertes Funktional von beschränktem Betrag. Wir nennen eine analytische Funktion $\varphi(w)$, die mit einer in $\mathfrak{G}$ regulären, schlichten nach (B) normierten Funktion $w=f(z)$ wieder eine derartige Funktion $\varphi(f(z))$ ergibt, so daß in

$$|F(\varphi(f))| > |F(f)|$$

gilt, eine „F-Schmiegungsfunktion von $\mathfrak{G}$".

Hat man eine Folge solcher Schmiegungsfunktionen φ_ν von $\mathfrak{G}$, so ist hiernach auch

$$\Phi_\nu(z) = \varphi_\nu(\varphi_{\nu-1}(\ldots(\varphi_0(z))\ldots))$$

eine in $\mathfrak{G}$ reguläre, schlichte, nach (B)-normierte Funktion und F-Schmiegungsfunktion von $\mathfrak{G}$.

$\mathfrak{G}$ soll nun auf ein schlichtes Normalgebiet vom Typus $\mathfrak{N}$ unter Erhaltung der Normierungsbedingungen abgebildet werden.

Def. 2. Wir sagen, F ist ein „den Typus $\mathfrak{N}$ erzeugendes Funktional" der Funktionenklasse $\mathfrak{F}$, wenn es, solange das durch eine Funktion $f(z)$ aus $\mathfrak{G}$ erzeugte Gebiet $\mathfrak{G}'$ noch nicht vom Typus $\mathfrak{N}$ ist, stets eine F-Schmiegungsfunktion von $\mathfrak{G}$, φ, gibt, so daß

$$|F(\varphi(f))| > |F(f)|$$

ist.

Es gilt dann der folgende

Satz. Hat man ein für die in $\mathfrak{G}$ regulären, schlichten, nach (B) normierten Funktionen definiertes, beschränktes, den Typus $\mathfrak{N}$ erzeugendes Funktional F mit Schmiegungsfunktionen φ_ν, so gilt:

1. Es existiert eine in $\mathfrak{G}$ reguläre, schlichte Funktion $f^*(z)$, für die $|F(f(z))|$ ein Maximum M annimmt. Diese Funktion leistet die Abbildung von $\mathfrak{G}$ auf ein Gebiet vom Typus $\mathfrak{N}$.

2. Hat man *irgendeine* Folge von F-Schmiegungsfunktionen φ_ν von $\mathfrak{G}$ eines den Typ $\mathfrak{N}$ erzeugenden Funktionals, für die $|F(\Phi_\nu)|$ mit $\nu \to \infty$ gegen M und $\Phi_\nu(z)$ gegen eine in $\mathfrak{G}$ reguläre Funktion $\Phi(z)$ konvergiert, so bildet $\Phi(z)$ das Gebiet $\mathfrak{G}$ auf ein schlichtes Gebiet vom Typus $\mathfrak{N}$ ab.

Der Beweis ergibt sich aus der Tatsache, daß die in $\mathfrak{G}$ regulären schlichten nach (B) normierten Funktionen $f(z)$ eine normale kompakte Familie $\mathfrak{F}$ bilden. Da der Betrag von F für diese Funktionen beschränkt ist, gibt es eine obere Grenze M von $|F(f)|$. Dann gibt es eine Funktion $f^*(z)$, so daß $|F(f^*)| = M$ ist oder man kann aus $\mathfrak{F}$ eine Folge von Funktionen $f_n(z)$ auswählen, für die $|F(f_n(z))|$ gegen M konvergiert, $\lim\limits_{n\to\infty} |F(f_n(z))| = M$. Aus dieser Funktionenfolge $f_n(z)$ läßt sich eine Teilfolge $f_{n_\nu}(z)$ auswählen, die gleichmäßig gegen eine Grenzfunktion $f^*(z) = \lim\limits_{\nu\to\infty} f_{n_\nu}(z)$ konvergiert. $f^*(z)$ kann nur eine schlichte in $\mathfrak{G}$ reguläre Funktion oder eine Konstante sein [9]. Letzteres scheidet wegen der Normierungsbedingungen aus. $f^*(z)$ ist selbst (B)-normiert und gehört somit zu $\mathfrak{F}$. Ferner ist $\lim\limits_{\nu\to\infty} |F(f_{n_\nu}(z))| = M$. Daher ist für jedes beliebig klein gewählte $\varepsilon > 0$ stets eine Schranke $N(\varepsilon)$ angebbar, so daß für $\nu > N(\varepsilon)$ stets $|F(f_{n_\nu})| > M - \varepsilon$ ausfällt. Es kann somit nur $|F(f^*(z))| \geq M$ sein. $|F(f^*(z))| > M$ scheidet jedoch aus, da $f^*(z)$ zu $\mathfrak{F}$ gehört und M die obere Grenze des Betrags $|F(f(z))|$ in $\mathfrak{F}$ ist. Somit[1]) ist $|F(f^*(z))| = M$. $f^*(z)$ leistet die Abbildung von $\mathfrak{G}$ auf ein Gebiet vom Typus $\mathfrak{N}$. Anderenfalls könnte man eine F-Schmiegungsfunktion φ angeben. Dann wäre aber $\varphi(f^*(z))$ eine Funktion von $\mathfrak{F}$ mit $|F(\varphi(f^*(z)))| > |F(f^*(z))| = M$. Widerspruch. In gleicher Weise folgt, daß die aus speziellen F-Schmiegungsfunktionen von $\mathfrak{G}$ als Grenzfunktion erhaltene in $\mathfrak{G}$ reguläre Funktion Φ zu $\mathfrak{F}$ gehört und wegen $|F(\Phi)| = M$ eine Abbildung von $\mathfrak{G}$ auf den Typ $\mathfrak{N}$ erzeugt.

Erzeugende Funktionale liefern die bekannten Extremaleigenschaften der Abbildungsfunktionen auf Normalgebiete $\mathfrak{N}$. Ist nämlich F ein für die Familie $\mathfrak{F}$ der in einem n-fach zusammenhängenden nach (B)-normierten Gebiet $\mathfrak{G}$ regulären, schlichten, nach (B) normierten Funktionen definiertes Funktional, dessen Betrag für eine Abbildung von $\mathfrak{G}$ auf ein Gebiet vom Typus $\mathfrak{N}$ und nur hierfür ein Maximum M annimmt, so ist es ein aus $\mathfrak{G}$ den Typus $\mathfrak{N}$ erzeugendes Funktional. Bildet nämlich $f^*(z)$ $\mathfrak{G}$ nach (B) normiert auf ein Gebiet vom Typ $\mathfrak{N}$ ab, dann ist nach Voraussetzung $|F(f^*(z))| = M$ und für eine Funktion $f_1(z)$ von $\mathfrak{F}$, die $\mathfrak{G}$ auf ein vom Typus $\mathfrak{N}$ verschiedenes Gebiet $\mathfrak{G}_1$ abbildet, gilt $|F(f_1(z))| < M$. Andererseits ist auch die Umkehrfunktion $z = f_1^{-1}(w)$ eine nach (B) normierte Funktion, die $\mathfrak{G}_1$ auf $\mathfrak{G}$ abbildet. $\varphi(w) = f^*(f_1^{-1}(w))$ ist dann eine nach (B) normierte Funktion, die $\mathfrak{G}_1$ auf ein Gebiet vom Typus $\mathfrak{N}$ abbildet. Wegen $\varphi(f_1(z)) = f^*(z)$ gilt dann $|F(\varphi(f_1(z)))| > |F(f_1(z))|$, w. z. b. w.

Für die praktische Brauchbarkeit eines Schmiegungsverfahrens ist jedoch die Art der Schmiegungsfunktionen ausschlaggebend. Diese müssen, da man

[1]) Falls F ein stetiges Funktional ist, folgt diese Gleichung unmittelbar aus

$$\lim_{\nu\to\infty} F(f_{n_\nu}(z)) = F\Big(\lim_{\nu\to\infty} f_{n_\nu}(z)\Big).$$

in einem praktischen Fall derartige Schmiegungsoperationen mehrfach anzuwenden hat, so beschaffen sein, daß man die Abbildungen $z_{\nu+1} = \varphi_\nu(z_\nu)$ numerisch oder graphisch bequem durchführen und durch Wahl freier Parameter den Besonderheiten der abzubildenden Zwischengebiete $\mathfrak{G}_\nu$ gut anpassen kann.

2.

Praktisch brauchbare Schmiegungsfunktionen kann man bis jetzt für ein- und zweifach zusammenhängende Gebiete angeben.

a) Bei der Abbildung eines einfach zusammenhängenden Teilgebietes $\mathfrak{G}$ des Einheitskreises mit dem Nullpunkt als innerem Punkt auf das Innere des Einheitskreises um 0 besteht die Normierung (B) darin, daß jedes derartige Teilgebiet durch die Abbildungsfunktion wieder auf ein Teilgebiet des Einheitskreises unter Festhaltung des Nullpunktes und gegebenenfalls des positiven Richtungselementes abgebildet werden soll. Als ein beschränktes, den Einheitskreis erzeugendes Funktional kann der Minimalabstand der Randpunkte des durch die Abbildungsfunktion erhaltenen Teilgebietes vom Nullpunkt angesehen werden. F-Schmiegungsfunktionen sind dann z.B. die bekannten KOEBE-CARATHÉODORYschen Wurzelabbildungen. Die Existenz weiterer praktisch brauchbarer F-Schmiegungsfunktionen folgt aus dem SCHWARZschen Lemma, wenn man das jeweils abzubildende Gebiet in ein Teilgebiet des Einheitskreises einbetten kann, das sich auf den ganzen Einheitskreis abbilden läßt [*4*]. Der Fall, daß das abzubildende Gebiet ∞ als inneren, 0 als äußeren Punkt enthält und die Abbildung auf das Äußere des Einheitskreises zu leisten ist, läßt sich mittels $Z = 1/z$ auf den obigen Fall zurückführen, doch hat man im reziproken Absolutbetrag des Maximalabstandes der Randpunkte der Bildgebiete von 0 direkt ein erzeugendes Funktional für diese Abbildungsaufgabe. Die Konstruktion der zugehörigen Schmiegungsfunktionen ist evident. Dieselben Funktionale und Schmiegungsfunktionen sind ferner in sinngemäßer Abänderung für die Abbildung auf das Innere bzw. Äußere eines Kreises um 0 mit einem von 1 verschiedenen Radius zu verwenden.

b) Man kann im vorigen Fall auch mit einer anderen Normierung (B) als der Nullpunktsnormierung auskommen. Hierbei betrachtet man das Innere des Einheitskreises als nichteuklidische Ebene mit POINCARÉscher Metrik. Nach der PICKschen Verallgemeinerung des SCHWARZschen Lemmas [*10*] wird durch jede im Einheitskreis $E_w(|w|<1)$ unimodular beschränkte reguläre Funktion das Innere des Einheitskreises auf sich oder einen Teil davon so abgebildet, daß der n.e. Abstand zweier Punkte nicht zunimmt. Wird daher ein Teilgebiet des Einheitskreises der z-Ebene durch eine analytische Funktion auf den schlichten Einheitskreis E_w abgebildet, so wird hierdurch ein in diesem Teilgebiet liegendes Gebiet n.e. expandiert. Für die n.e. Entfernung zweier entsprechender Punktepaare gilt dann die Ungleichung $E(w_1, w_2) > E(z_1, z_2)$ mit $E(z_1, z_2) = 2\,\mathfrak{Ar}\,\mathfrak{T}\mathfrak{g}\,\psi(z_1, z_2)$, $\psi(z_1, z_2) = \frac{|z_1 - z_2|}{1 - |z_1 z_2|}$. Daraus folgt

für $z_1 \to z_2$ die Ungleichung

$$|F(w)| = \frac{\left|\frac{dw}{dz}\right|}{1-|w|^2} > \frac{1}{1-|z|^2}.$$

$|F(w)|$ ist ein Maß für die Größe der n.e. Expansion an der Stelle z. Es ist für alle Funktionen, die ein Teilgebiet $\mathfrak{G}$ des Einheitskreises auf ein ebensolches abbilden in einem inneren Punkt z_0 am größten bei den Funktionen, die $\mathfrak{G}$ auf den vollen Einheitskreis abbilden und kann daher als ein den Einheitskreis erzeugendes Funktional angesehen werden. Es ist wegen der bekannten Verzerrungssätze beschränkt und sogar stetig. Die zugehörigen Schmiegungsfunktionen sind in gleicher Weise wie unter a) zu erhalten, es kann jedoch die Nullpunktnormierung wegbleiben. Die Abbildung von $\mathfrak{G}$ auf den Einheitskreis kann nachträglich durch die Zuordnung zweier Linienelemente eindeutig festgelegt werden[2]).

c) Für die konforme Abbildung eines zweifach zusammenhängenden schlichten Gebietes $\mathfrak{G}$, bei dem die Randkontinuen jeweils mindestens zwei Punkte haben, kann man o. B. d. A. als Bedingungen (B) annehmen, daß $\mathfrak{G}$ und die hieraus durch Abbildung entstehenden Gebiete innerhalb eines konzentrischen Kreisringes um 0 $(0<r<|z|<R)$ liegen und die Ränder dieser Gebiete den Kreis $|z|=r$ ganz umschließen. Ist D der Maximalabstand der Randpunkte der äußeren Begrenzung von $\mathfrak{G}$ von 0, d der Minimalabstand der inneren Randpunkte von 0, so setzen wir $r=d$ und $R=D$. Allgemein sei D_ν bzw. d_ν der Maximal- bzw. Minimalabstand der Randpunkte der äußeren bzw. inneren Begrenzung eines aus $\mathfrak{G}$ durch eine nach (B) normierte Funktion $f(z)$ erhaltenen Gebietes $\mathfrak{G}_\nu$. Als ein den Kreisring erzeugendes, beschränktes Funktional kann hier die Größe $\frac{d_\nu}{D_\nu} = F(f_\nu(z))$ Verwendung finden, der reziproke Wert des äußeren Moduls nach ALBRECHT [5][3]). Schmiegungsfunktion ist hier jede nullpunktnormierte Schmiegungsfunktion nach a) für die Abbildung des einfach zusammenhängenden endlichen Gebietes, das vom äußeren Rand des Ringgebietes $\mathfrak{G}_\nu$ begrenzt wird, auf den Kreis $|z_\nu|<D_\nu$, sowie jede den Punkt ∞ festhaltende Schmiegungsfunktion nach a) für die Abbildung des einfach zusammenhängenden unendlichen Gebietes, das vom inneren Rand des Ringgebietes $\mathfrak{G}_\nu$ begrenzt wird auf $|z_\nu|>d$. Ist nämlich nach Ausübung einer solchen Abbildung $d_{\nu+1}$ bzw. $D_{\nu+1}$ der neue Minimal- bzw. Maximalabstand der inneren bzw. äußeren Begrenzung von 0, so gilt bei den

[2]) Nach UNKELBACH [11] läßt sich jede mit $f(0)=0$ und $f'(0)>0$ normierte unimodular beschränkte schlichte Funktion unter gewissen Bedingungen durch kongruente Abbildungen des Einheitskreises approximieren. Aus b) ergibt sich leicht ein entsprechender Sachverhalt bei n.e. Kongruenz.

[3]) In dieser Arbeit von ALBRECHT wird beim Existenzbeweis für die Abbildungsfunktion S. 172 angenommen und benutzt, daß der äußere Modul eines Ringgebietes ein stetiges Funktional ist. Das ist jedoch, wie Herr GAIER in einem Brief an den Verf. bemerkt, nicht der Fall. Wie aus 1. folgt, genügt jedoch für den Existenzbeweis schon die Tatsache, daß der äußere Modul bzw. sein reziproker Wert ein beschränktes Funktional ist.

Abbildungen der ersten Art $d_{\nu+1} > d_\nu$, $D_{\nu+1} = D_\nu$, bei den der zweiten Art $d_{\nu+1} = d_\nu$, $D_{\nu+1} < D_\nu$. In jedem Falle ist also $\frac{d_{\nu+1}}{D_{\nu+1}} > \frac{d_\nu}{D_\nu}$. Dieses Verfahren läßt sich solange fortsetzen als es noch Punkte in $d_{\nu+1} < |z_{\nu+1}| < D_{\nu+1}$ gibt, die nicht zu $\mathfrak{G}_{\nu+1}$ gehören.

d) Bei einem Iterationsverfahren von GRÖTZSCH [*12*] für die Abbildung n-fach zusammenhängender Gebiete etwa auf die von konzentrischen Kreisbogenschlitzen um 0 begrenzte Ebene wird jeweils ein Randkontinuum auf dem die Schwankung von $\log|z|$ am größten ist, auf einen Kreisbogenschlitz mit 0 als Mittelpunkt abgebildet. Dabei wird bei jeder Iteration die Kreisbogenform eines vorher transformierten Randkontinuums wieder zerstört, ähnlich wie bei den Verfahren von KOMATU [*13*] und GRAESER [*14*] für die Abbildung von Ringgebieten. Jede Abbildung auf einen Kreisbogenschlitz um 0 kann aber selbst nur iterativ erfolgen. Es bleibt somit für $n \geq 3$ die Aufgabe ein Schmiegungsverfahren zu finden, das ähnlich wie das unter c) angegebene modifizierte ALBRECHTsche Verfahren in einer einfachen Folge von praktisch durchführbaren Schmiegungsschritten direkt zum Ziele führt.

Literatur

[*1*] KOEBE, P.: Abhandlungen zur Theorie der konformen Abbildung I. J. reine angew. Math. **145**, 177—195 (1915). — [*2*] RINGLEB: Numerische und graphische Verfahren der konformen Abbildung. Heidelberg 1939. — [*3*] HEINHOLD, J.: Ein Schmiegungsverfahren der konformen Abbildung. Sitzgsber. bayer. Akad. Wiss., Math.-naturw. Kl. **1948**, 203—222. — [*4*] HEINHOLD, J., u. R. ALBRECHT: Zur Praxis der konformen Abbildung. Rendiconti del Circolo matem. Palermo **3**, 2—19 (1954). — [*5*] ALBRECHT, R.: Zum Schmiegungsverfahren der konformen Abbildung. Z. angew. Math. u. Mechanik **32**, 316—318 (1952). — [*6*] ALBRECHT, R.: Iterationsverfahren zur konformen Abbildung eines Ringgebietes auf einen konzentrischen Kreisring. Sitzgsber. bayer. Akad. Wiss., Math.-naturw. Kl. **1954**, 169—178. — [*7*] KOEBE, P.: Abhandlungen zur Theorie der konformen Abbildung. IV. Abbildung mehrfach zusammenhängender schlichter Bereiche auf Schlitzbereiche. Acta Math. **41**, 305—344 (1918). — [*8*] NEHARI, Z.: Conformal Mapping, S. 341 ff. New York 1952. — [*9*] CARATHÉODORY, C.: Funktionentheorie, Bd. I, S. 193. Basel 1950. — [*10*] CARATHÉODORY, C.: Funktionentheorie, Bd. II, S. 15. Basel 1950. — [*11*] UNKELBACH, H.: Über die Approximation schlichter konformer Abbildungen durch kongruente Iterationen spezieller sterniger oder konvexer konformer Abbildungen. Math. Z. **58**, 63—70 (1953). — [*12*] GRÖTZSCH, H.: Zur konformen Abbildung mehrfach zusammenhängender schlichter Bereiche. Ber. sächs. Akad. Wiss. Leipzig, Math.-phys. Kl. **83**, 67—76 (1931). — [*13*] KOMATU, Y.: Ein alternierendes Approximationsverfahren für konforme Abbildung von einem Ringgebiet auf einen Kreisring. Proc. Jap. Akad. **21**, 146—155 (1945). — [*14*] GRAESER, E.: Über konforme Abbildung des allgemeinen zweifach zusammenhängenden schlichten Bereiches auf die Fläche des Einheitskreises. Diss. Leipzig 1930.

Technische Hochschule München, Lehrstuhl für angewandte Mathematik und mathematische Statistik

(Eingegangen am 28. Juli 1956)

HOFMANN, J. E.
Math. Zeitschr. Bd. 67, S. 139—146 (1957)

Zur Entwicklungsgeschichte der EULERschen Summenformel

Herrn G. FABER zum 80. Geburtstag am 5. April 1957 gewidmet

Von

JOS. E. HOFMANN

1. Im mathematischen Gesamtwerk EULERs nehmen die Beiträge zur Reihenlehre eine interessante Sonderstellung ein. Auf diesem Gebiete ist EULER stärker Autodidakt denn auf anderen. Erst in Petersburg gewinnt er im Ringen um die Ergebnisse der WALLISschen *Arithmetica infinitorum*[1]) von 1656 — eines damals bereits völlig veralteten Werkes — neue Gesichtspunkte. In selbständig gestaltender Weiterbildung von mancherlei Anregungen, die ihm über die Kollegen aus dem engen Kreise der Akademie zufließen, sind die ersten Beiträge zur Reihenlehre entstanden, im Anfang noch ungelenk in Bezeichnung und Ausdrucksweise, später flüssiger und lebendiger, bis sich auch hier die Meisterhand des großen Genies bekundet. Wir verdanken den Herren C. BOEHM und G. FABER die ausgezeichnete dreibändige Ausgabe der Einzelschriften EULERs zur Reihenlehre[2]) und Herrn FABER eine wertvolle Zusammenfassung der abgedruckten Abhandlungen nach sachlichen Gesichtspunkten[3]). Erst auf dieser Unterlage bauend, ist es möglich, auch den historischen Gesichtspunkt stärker in den Vordergrund treten zu lassen. Im folgenden soll ein Hauptthema der Frühzeit der EULERschen Reihenlehre, nämlich die berühmte Summenformel, in ihrer Entwicklungsgeschichte etwas näher verfolgt werden. Hierbei kann ich mich dank des liebenswürdigen Entgegenkommens von Herrn O. SPIESS-Basel, dem Herausgeber der BERNOULLI-Ausgabe, und von Herrn P. JUSCHKEWITSCH-Moskau, dem Sekretär der EULER-Kommission der Moskauer Akademie der Wissenschaften, auf noch unediertes Material stützen, dessen Existenz erst seit kurzem bekannt geworden ist. Ich möchte nicht versäumen, diesen beiden Herren, sowie Herrn K.-R. BIERMANN-Berlin, der die Verbindung zu Herrn JUSCHKEWITSCH übernommen hat, und Herrn K. HAAS-Heidelberg, der mir wichtige Auszüge aus schwer zugänglichen älteren Werken zur Verfügung gestellt hat, meinen tiefempfundenen Dank für die empfangene Unterstützung auszusprechen.

[1]) WALLIS, J.: *Arithmetica infinitorum*, Oxford „1655" = *Opera mathematica* I, Oxford 1695. Das Ausgabejahr des Erstdruckes ist nach der englischen Jahreszählung gerechnet, die den 25. 3. als Jahresanfang nimmt. Nach der festländischen Jahreszählung mit dem Jahreswechsel am 1. 1. ist das Ausgabejahr 1656. EULER dürfte nicht den verhältnismäßig seltenen Erstdruck verwendet haben, sondern den viel häufigeren Nachdruck.

[2]) EULER, L.: *Opera mathematica* (1) XIV/XVI: *Commentationes analyticae ad theoriam serierum pertinentes*, Leipzig-Berlin-Zürich-Basel 1924/35, im folgenden zitiert als *EO* mit Bandnummer und nachfolgender Seitenzahl.

[3]) *EO* XVI*, S. VII/CXII.

2. Die erste Abhandlung EULERs zur Reihenlehre[4]), im ENESTRÖMschen Schriftenverzeichnis[5]) als Nummer 19 = [E 19] bezeichnet, enthält noch nicht viel Bedeutendes, wohl aber den weittragenden und dann von EULER konsequent zur Anwendung gebrachten Gedanken, Reihenglieder durch bestimmte Integrale auszudrücken und Reihensummen durch bestimmte Integrale in „geschlossener Form" darzustellen[6]). Diese Abhandlung ist in Weiterbildung einer Studie Chr. GOLDBACHs über rekurrente Reihen[7]) entstanden. Kennzeichnend ist etwa die Definition des natürlichen Logarithmus aus $\frac{1-x^0}{0} = -\ln x$, die wahrscheinlich nicht selbständig[8]) ist.

Aus E 20, der zweiten Abhandlung EULERs zur Reihenlehre[9]), ist die Art bemerkenswert, wie EULER die Summe $s_2 = \sum_{\nu=1}^{\infty} \frac{1}{\nu^2}$ aus $\int_0^1 \frac{\ln(1-t)\cdot dt}{-t}$ erklärt — hierbei vielleicht einer mündlichen Anregung aus Joh. BERNOULLIs Unterweisungen in Basel folgend[10]). Durch Zerspalten des Integrals in die

[4]) *De progressionibus transcendentibus, seu quarum termini generales algebraice dari nequeunt: Commentarii academiae scientiarum Petropolitanae* [= *CP*] **5** (1730/31), 1738 = *EO* XIV, S. 1/24, der Akademie am 28. 11. 1729 a. St. vorgelegt.

[5]) ENESTRÖM, G.: *Verzeichnis der Schriften Leonhard Eulers*, Leipzig 1910/13 = Jahresbericht der Deutschen Mathematiker-Vereinigung, Erg.-Bd. IV.

[6]) Dan. BERNOULLI berichtet übrigens im Brief vom 7. 3. 1742 an EULER, er habe sich schon vor seiner Wirksamkeit in Petersburg (1725/33) ebenfalls mit der Herstellung summierbarer Reihen (in geschlossener Form) unter Verwendung von Differentiationen und Integrationen beschäftigt. Im nächsten Brief an EULER vom 14. 4. 1742 gesteht er allerdings, nachdem er die unten zu erwähnende Abhandlung[15]) E 41 aus *CP* **7** (1734/35), 1740 gesehen hatte: „Ich sehe nun freylich, daß Sie die differentiationes und integrationes serierum anders nehmen, und zwar auf eine viel nützlichere und ingeniosere Weise als ich." ... Die erwähnten Briefstellen finden sich in P.-H. FUSS: *Correspondance mathématique et physique de quelques célèbres géomètres du XVIIIème siècle* II, Petersburg 1843, S. 487/88 und 493.

[7]) GOLDBACH, Chr.: *De terminis generalibus serierum*, *CP* **3** (1728), 1732, S. 164/73.

[8]) Die gleichwertige Definition $\ln(1+t) = \lim_{n\to\infty} \frac{(1+t)^{1/n}-1}{1/n}$ findet sich in E. HALLEY: *A most compendious and facile Method for constructing the Logarithmes ...*, *Philosophical Transactions* [= *PT*] **19**, Nr. 216 für III/V 1695, S. 88/97. Dort wird erstmals der (allerdings nur formal geglückte) Versuch unternommen, die logarithmische Reihe als Grenzfall der binomischen Reihe anzusehen und hieraus den Logarithmus auf algorithmischem Wege zu erklären. Es ist kaum wahrscheinlich, daß EULER direkt an diesen Aufsatz anschließen konnte; die unmittelbare Vorlage ist nicht feststellbar.

[9]) *De summatione innumerabilium progressionum*, *CP* **5** (1730/31), 1738 = *EO* XIV, S. 25/41, der Akademie am 5. 3. 1731 a. St. vorgelegt.

[10]) Schon LEIBNIZ hatte im Brief vom 9. [19.] 11. 1696 (d.i. 9. 11. a. St. = 19. 11. n. St.) an Joh. BERNOULLI festgestellt, daß $\sum_{\nu=1}^{\infty} \frac{1}{\nu^2} = \int_0^1 \frac{\ln(1-t)\cdot dt}{-t}$ und $\sum_{\nu=1}^{\infty} \frac{(-1)^{\nu+1}}{\nu^2} = \int_0^1 \frac{\ln(1+t)\cdot dt}{t}$ ist. Vgl. LEIBNIZens *Mathematische Schriften*, ed. C. I. GERHARDT, Bd. III, Halle 1855, S. 337. Fast gleichzeitig und unabhängig von LEIBNIZ kam Jak. BERNOULLI zur nämlichen Integraldarstellung für die beiden Summen: *prop.* 44 der

beiden Bestandteile

$$\int_0^x \cdots = x + \frac{x^2}{2^2} + \frac{x^3}{3^2} + \cdots \quad \text{und} \quad \int_x^1 \cdots = \ln x \cdot \ln y + \int_0^y \frac{\ln(1-t)\cdot dt}{-t}$$

mit $x+y=1$ führt er die numerische Bestimmung von s_2 mit $x=y=\frac{1}{2}$ zurück auf

$$1 + \frac{1}{2\cdot 2^2} + \frac{1}{2^2\cdot 3^2} + \frac{1}{2^3\cdot 4^2} + \cdots + (\ln 2)^2.$$

Euler interessiert sich — wohl im Zusammenhang mit ähnlichen Versuchen[11]) Goldbachs und Dan. Bernoullis, über die er nur andeutungsweise unter-

---

dritten Reihendissertation vom 14. [24.] 11. 1696, Wiederdruck im Anhang zur nachgelassenen *Ars conjectandi*, die 1713 zu Basel unter Mitwirkung von Nikl. I. Bernoulli gedruckt wurde.

[11]) Schon 1723 beginnt die Diskussion zwischen den beiden Freunden um den (geschlossenen) Summenwert von s_2. Dan. Bernoulli verweist im Brief vom 2. 10. 1723 [Fuss[6]) II, S. 180/81] auf Jak. Bernoullis erste Reihendissertation vom 7. [17.] 6. 1689, woselbst der Autor in einer Ergänzung zu *prop.* 17 bemerke, er habe trotz erheblicher Anstrengungen s_2 nicht bestimmen können. In seiner Antwort vom 4. 11. 1723 berichtet Goldbach über seine ersten Reihenstudien und erwähnt, er besitze die *Ars conjectandi*[10]), habe aber früher bei der Durcharbeit große Schwierigkeiten gehabt. Er halte die Summierung von s_2 für nicht einfacher wie die Kreisquadratur (Fuss, S. 185).

Das ungelöste Problem erscheint wieder im Jahre 1728. Jetzt steht die numerische Approximation im Vordergrund. Am 29. 8. [9. 9.] 1728 schreibt Dan. Bernoulli, er habe eine kurze Methode zur Berechnung von s_2 gefunden, aus der $s_2 \approx \frac{8}{5}$ folge (Fuss, S. 263/64). Später muß er angedeutet haben, daß $s_2 - 1 = \sum_{\nu=1}^{\infty} \frac{1}{\nu^2(\nu+1)}$ ist (erkennbar aus Goldbachs Schreiben vom 2. 10. 1730: Fuss, S. 395). Den Beweis führt er im Brief vom 8. [19.] 10. 1730 (Fuss, S. 398) auf Grund der Umformung

$$\sum_{\nu=1}^{\infty} \left\{ \frac{1}{(\nu+1)^2} - \frac{1}{\nu^2(\nu+1)} \right\} = \sum_{\nu=1}^{\infty} \left\{ \frac{\nu}{(\nu+1)^2} - \frac{\nu-1}{\nu^2} \right\} = 0.$$

Die einfachere Umformung

$$\sum_{\nu=1}^{\infty} \left\{ \frac{1}{\nu^2} - \frac{1}{\nu^2(\nu+1)} \right\} = \sum_{\nu=1}^{\infty} \left\{ \frac{1}{\nu} - \frac{1}{\nu+1} \right\} = 1$$

scheint ihm entgangen zu sein. Die Näherung $\frac{8}{5}$ folgt durch Berücksichtigung der ersten drei Glieder der neuen Reihe:

$$s_2 > 1 + \frac{1}{1^2\cdot 2} + \frac{1}{2^2\cdot 3} + \frac{1}{3^2\cdot 4} = \frac{29}{18} > \frac{8}{5}.$$

Dazwischen schieben sich die Verfahren Goldbachs. Dieser verwendet im Schreiben vom 31. 1. 1729 (Fuss, S. 281/82) die Ungleichungen

$$\frac{256}{(16\nu-7)(16\nu+9)} < \frac{1}{\nu^2} < \frac{4}{(2\nu-1)(2\nu+1)} \quad \text{für} \quad \nu > 1$$

richtet worden sein mag — auch für die reine Zahlenrechnung. Zur Berechnung von ln 2 bedient er sich nicht der zweckmäßigsten Form aus $\ln \frac{3+1}{3-1}$, die er ersichtlich noch nicht kennt[12]), sondern der aus $\ln \frac{1}{1-\frac{1}{2}}$ folgenden, die er aus Wallis haben dürfte[13]). Vielleicht ist Euler bei dieser Gelegenheit auf den Gedanken gekommen, es an Stelle von $(\ln 2)^2$ mit einer Näherung zu versuchen, die in etwa mit dem in seinem Sinne zu $(\ln 2)^2$ gleichwertigen Ausdruck π^2 zusammenhängt[14]), und mag sich auf diesem Wege den Zugang zu der berühmten Darstellung von s_2 durch $\frac{1}{6}\pi^2$ gebahnt haben[15]).

3. Die numerische Berechnung von Reihensummen, deren ν-tes Glied $t(\nu)$ als geschlossener Ausdruck von ν angesehen werden kann: das ist zunächst Eulers Ziel. Auf diesem Gebiet gelingt ihm dann der erste große Erfolg, nämlich die Aufstellung der Summenformel[16]) in E 25. Sie erscheint dort in der Form $s = \int t \cdot dn + \alpha t + \beta \frac{dt}{dn} + \gamma \frac{d^2t}{dn^2} + \cdots$.

Hier ist unter s die Summe der ersten n Glieder dieser Reihe zu verstehen. Die konstanten Zahlenkoeffizienten werden aus den Rekursionsformeln $\alpha = \frac{1}{2}$, $\beta = \frac{1}{2}\alpha - \frac{1}{6}$, $\gamma = \frac{1}{2}\beta - \frac{1}{6}\alpha + \frac{1}{24}$ usw. festgestellt. Euler ist sich übrigens zu diesem Augenblick noch keineswegs über die Tragweite seiner Formel im klaren; er glaubt eine spezielle Summenformel vor sich zu haben und

---

und findet durch Summieren

$$\frac{16}{16n-7} < \sum_{\nu=n}^{\infty} < \frac{2}{2n-1}.$$

Er berechnet zuerst die ersten 6, dann die ersten 11 Glieder von s_2 und schätzt die Reste vermittels der soeben angegebenen Schranken ab. Die sich ergebenden Näherungswerte unterscheiden sich im zweiten Fall nur mehr um drei Einheiten der 4. Dezimale.

Dazu tritt Goldbachs Bemerkung, $s_2 - 1$ lasse sich auch in der Form

$$\frac{1}{2^2-1} + \frac{1}{3^2-1} + 0 + \frac{1}{5^2-1} + \frac{1}{6^2-1} + \frac{1}{7^2-1} + 0 + 0 + \frac{1}{10^2-1} + \cdots$$

darstellen (Fuss, S. 283). Dan. Bernoulli hält diese Entwicklung nicht für richtig. Nach längerer Diskussion deutet Goldbach in einem Brief vom April 1729 das einzuschlagende Beweisverfahren an (Fuss, S. 296). Einen Beweis unter Bezugnahme auf Goldbachs Reihe veröffentlichte Euler in E 72: *Variae observationes circa series infinitas*, *CP* **9** (1737), 1744 = *EO* XIV, S. 216/44, insbesondere S. 226/27.

[12]) Die Berechnung natürlicher Logarithmen in der Form $\ln \frac{a+x}{a-x}$ erscheint erstmals im Druck in J. Gregory: *Exercitationes geometricae*, London 1668, S. 9/13 und ist von hier in Halleys Aufsatz [8]) übergegangen.

[13]) Es handelt sich um die Note in den *PT* **3**, Nr. 38 vom 17. [27.] 8. 1668, S. 753/56, die auch in Wallis' *Opera*[1]) übergegangen ist.

[14]) Entsprechung der logarithmischen Reihe mit jener für den Arcus Tangens.

[15]) Die erste Darstellung findet sich in der Abhandlung E 41: *De summis serierum reciprocarum*, *CP* **7** (1734/35), 1740 = *EO* XIV, S. 73/86, der Petersburger Akademie am 5. 12. 1735 a. St. vorgelegt, jedoch nicht abgeliefert.

[16]) *Methodus generalis summandi progressiones*, *CP* **6** (1732/33), 1738 = *EO* XIV, S. 42/72, der Akademie am 30. 6. 1732 a. St. vorgelegt.

$t(\nu)$ als ganzrationale Funktion von ν ansehen zu müssen. Deshalb beschränkt er sich auf das triviale Beispiel $t = \nu^2 + 2\nu$ und fährt dann mit einer großen Zahl anderer spezieller Summenverfahren fort, die heute kaum mehr interessieren. Eines davon bezieht sich auch auf die Teilreihen der harmonischen Reihe und führt nahe an den Begriff der EULERschen Konstanten heran.

4. Diese Konstante wird erstmals genauer in der Abhandlung[17]) E 43 bestimmt. Hier geht EULER aus von einer Teilreihe der harmonischen Reihe, schätzt deren Summe arithmetisch unter Verwendung des mittleren Gliedes bzw. des arithmetischen Mittels der äußeren Glieder ab, vergleicht die Reihensumme mit dem natürlichen Logarithmus des Quotienten aus dem Anfangsglied und dem Schlußglied und findet schließlich[18])

$$\ln p = \lim_{n\to\infty} \sum_{\nu=n+1}^{np} \frac{1}{\nu}.$$

Daraus erhält er dann durch kühne Schlüsse (formales Rechnen mit bedingt konvergenten Reihen) Darstellungen für ln 2, ln 3, ln 4, ln 5, ln 6, ... Hierauf summiert er

$$\ln \frac{\nu+1}{\nu} = \frac{1}{\nu} - \frac{1}{2\nu^2} + \frac{1}{3\nu^3} \mp \cdots$$

und erhält

$$\sum_{\nu=1}^{n} \frac{1}{\nu} - \ln(n+1) = \frac{1}{2}\sum_{\nu=1}^{n} \frac{1}{\nu^2} - \frac{1}{3}\sum_{\nu=1}^{n} \frac{1}{\nu^3} \pm \cdots.$$

Ohne die Einzelrechnung auszuführen, stellt er fest, daß das rechts stehende Aggregat mit $n \to \infty$ die Summe $C = 0{,}577218$ hat.

[17]) *De progressionibus harmonicis observationes*, *CP* **7** (1734/35), 1740 = *EO* XIV, S. 87/100, der Akademie am 11. 3. 1734 a. St. vorgelegt.

[18]) EULER hatte keine Ahnung davon, daß schon in P. MENGOLI: *Geometria speciosa*, Bologna 1659, die auf der Treppenmethode beruhenden Ungleichungen

$$\frac{1}{n+1} + \frac{1}{n+2} + \cdots + \frac{1}{np} < \ln p < \frac{1}{n} + \frac{1}{n+1} + \cdots + \frac{1}{np-1}$$

erscheinen, aus denen die EULERsche Formel sogleich durch Grenzübergang hervorgeht. Wahrscheinlich hatte er aber erfahren (brieflicher Hinweis von Herrn O. SPIESS), daß Joh. II. BERNOULLI, der damals 19jährige Sohn Joh. I. und jüngere Brüder Dan. BERNOULLIs, die Teilreihen der harmonischen Reihe durch einen geschlossenen Ausdruck angenähert hatte. Über diese Angelegenheit, von der Dan. BERNOULLI im Frühjahr 1729 unterrichtet worden sein muß, wurde nämlich auch GOLDBACH informiert. Der fragliche Brief Dan. BERNOULLIs trägt nach FUSS[6]), S. 250 das Datum 20. 2. 1728 a. St. Hier muß jedoch ein Lesefehler bezüglich der Jahreszahl vorliegen. Das richtige Datum ist wohl der 3. 3. 1729 n. St. Der Brief bezieht sich nämlich auf Fragen, die erst im Jahr 1729 zwischen Dan. BERNOULLI und GOLDBACH verhandelt wurden.

Dan. BERNOULLI hat die erhaltenen Andeutungen aus Basel sofort verstanden, wie das aus dem Hinweis bei FUSS, S. 252/53 hervorgeht. Ein späterer Bericht Dan. BERNOULLIs an GOLDBACH vom 28. 4. [9. 5.] 1729 (FUSS, S. 300/01) bestätigt dies. Auch GOLDBACH merkt, daß es sich um die Verwendung von Logarithmen handeln müsse: Brief an Dan. BERNOULLI vom April 1729 (FUSS, S. 296/97).

5. In der Abhandlung[19]) E 46 wird das Summenproblem von ganz anderer Seite aus angegriffen. Euler geht aus von einer summierbaren (endlichen oder unendlichen) Reihe mit lauter positiven Gliedern, deren ν-tes Glied wiederum als angebbarer Ausdruck $t(\nu)$ hingeschrieben werden kann. Er zeichnet die durch $y=t(x)$ und $z=t(x+1)$ bestimmten Kurven und nähert die Summe $s=\sum_{\nu=1}^{n} t(\nu)$ unter Verwendung der Treppenregel durch die anschaulich gedeuteten Integrale $\int_0^n y\cdot d\nu$ und $\int_1^{n+1} z\cdot d\nu$ an[20]). Zuerst verbessert er die Schranken durch Ubergang zur Trapezmethode, dann unter Mitverwendung von Strecken, die sich nur wenig von den Tangenten an die beiden gezeichneten Kurven in den gegebenen Kurvenpunkten unterscheiden:

$$s=\int_1^{n+1} t(\nu)\,d\nu+\frac{t(1)-t(n)}{2}+\frac{t(1)-t(2)}{12}-\frac{t(n+1)-t(n+2)}{12}.$$

Daraus gewinnt er sehr gut brauchbare Näherungen für die Teilreihen der harmonischen Reihe. Zum Beispiel berechnet Euler die Summe der ersten 10 Glieder $\frac{1}{1}+\frac{1}{2}+\cdots+\frac{1}{10}$ direkt, setzt dann $t(\nu)=\frac{1}{n+\nu}$ und erhält mit $n=999990$ einen bis auf 2 Einheiten der 4. Dezimale richtigen Näherungswert[21]) für die Summe aus der ersten Million der Glieder der harmonischen Reihe. Das Ganze dient auch zu einer neuen und rechnerisch besser überblickbaren numerischen Bestimmung der Konstanten C. Ist die Reihe sogar absolut konvergent, dann kann Euler mit $n\to\infty$ die Glieder $t(n)$, $t(n+1)$, $t(n+2)$ in seiner Formel weglassen. Beispiel: er berechnet die ersten 10 Glieder der Summe s_2, erhält $\int_{11}^{\infty} t(\nu)\cdot d\nu=\frac{1}{11}$ und einen bis auf 2 Einheiten der 5. Dezimale richtigen Zahlenwert von s_2.

6. An diesen speziellen Ergebnissen erkennt Euler, daß seine Summenformel nicht nur auf Reihen beschränkt ist, deren Glieder $t(\nu)$ ganzrationale Funktionen der Gliederzahl ν sind. Sie ist auch dann verwendbar, wenn $t(\nu)$ als Taylor-Reihe dargestellt werden kann, also im Sinne der Eulerschen Auffassung allgemein. Das ist der wesentliche Inhalt der Abhandlung[22])

[19]) *Methodus universalis serierum convergentium summas quam proxime inveniendi*, *CP* **8** (1736), 1741 = *EO* XIV, S. 101/07, der Akademie am 9. 6. 1735 a. St. vorgelegt.

[20]) Stillschweigend wird vorausgesetzt, daß die Reihenglieder monoton abnehmen, und zwar so, daß auch $t(\nu+1)-t(\nu)>t(\nu)-t(\nu-1)$ ist.

[21]) Die im Brief an Joh. Bernoulli vom 20. 6. [1. 7.] 1740 gemachten Mitteilungen Eulers über seine neuen Ergebnisse — abgedruckt in G. Enestrom: *Der Briefwechsel zwischen L. Euler und Joh. I. Bernoulli*, Abt. III, *Bibl. math.* (3) **6**, 1904, S. 62/67 — führen zu Joh. Bernoullis Antwort vom 31. 8. 1740. Hier tut Joh. Bernoulli so, als habe er die Ergebnisse Eulers aus eigener Kraft nacherfunden [Fuss[6]), S. 45/47]. In Wirklichkeit hat er sich jedoch nur das Verfahren seines Sohnes Joh. II. angeeignet. Siehe oben[18]).

[22]) *Inventio summae cujusque seriei ex dato termino generali*, *CP* **8** (1736), 1741 = *EO* XIV, S. 108/23, der Akademie am 13. 10. 1735 a. St. vorgelegt.

E 47. Jetzt erst wird die Summenformel bewiesen, und zwar vermittels des TAYLORschen Lehrsatzes. Die Art, wie die nachfolgenden Summen $\sum_{\nu=1}^{n} \nu^k$ $(k = 1, 2, \ldots, 16)$ hingeschrieben werden, zeigt uns, daß EULER die *Ars conjectandi*[10]) Jak. BERNOULLIS noch immer nicht kennt[23]). Die weiteren Beispiele (EULERsche Konstante, Partialsummen der harmonischen Reihe, numerische Berechnung von s_2, s_3 und s_4) gehen nicht mehr über das nunmehr zu Erwartende hinaus.

7. In dieser Abhandlung bricht EULER die Summenformel mit dem Glied bei der 15. Ableitung ab, gibt jedoch 20 richtige Dezimalen von s_2. Daraus schließt Herr FABER[24]), bei der endgültigen Redaktion der Abhandlung E 47 müsse EULER bereits im Besitz des geschlossenen Ausdrucks für s_2 gewesen sein. Diese Ansicht kann ich nunmehr durch weiteres Material stützen.

In der nächstfolgenden Abhandlung[25]) E 55 wird die Summenformel bis zum Glied mit der 17. Ableitung fortgesetzt, aber gerade an dieser Stelle steht im Originaldruck eine unrichtige Zahl. Das könnte ein Druckfehler sein, kein Rechenfehler. Dem ist aber nicht so. Am 8. [19.] 6. 1736 hat nämlich EULER einen Brief an J. STIRLING geschrieben, auf dessen *Methodus differentialis*[26]) er durch die eingehende Anzeige in den *Nova Acta Eruditorum*[27]) aufmerksam geworden war und in der er dann bei genauer Durchsicht drei ausgezeichnete numerische Verfahren zur Bestimmung von s_2 vorgefunden hatte[28]). In diesem bisher noch unedierten Brief[29]) teilt EULER neben dem geschlossenen Ausdruck für s_2 und anderem auch die Summenformel mit,

[23]) Dort sind nämlich auf S. 96/98 nicht nur die Summen für $k = 1, 2, \ldots, 10$ angegeben, sondern auch das allgemeine Bildungsgesetz erscheint:

$$\sum_{\nu=1}^{n} \frac{1}{\nu^k} = \frac{n^{k+1}}{k+1} + \frac{n^k}{1 \cdot 2} + A \cdot \frac{k\, n^{k-1}}{2} + B \cdot \frac{k(k-1)(k-2)}{2 \cdot 3 \cdot 4}\, n^{k-3} + \cdots .$$

Daraus hätte EULER seine Summenformel unmittelbar ablesen können. Sicherlich hätte EULER schon damals auf Jak. BERNOULLI hingewiesen, wenn er die fragliche Stelle gekannt hätte. Später hat er das Versäumte ausdrücklich nachgeholt: Abhandlung E 393: *De summis serierum numeros Bernoullianos involventium*, *Novi CP* **14** (1769) I, 1770 = *EO* XV, S. 91/130. Die Abhandlung wurde der Petersburger Akademie am 18. 8. 1768 a. St. vorgelegt.

[24]) *EO* XVI*, S. XXII.

[25]) *Methodus universalis series summandi ulterius promota*, *CP* **8** (1736), 1741 = *EO* XIV, S. 124/37, der Akademie am 17. 9. 1736 vorgelegt.

[26]) J. STIRLING: *Methodus differentialis sive tractatus de summatione et interpolatione serierum infinitarum*, London 1730.

[27]) Novemberheft des Jahrganges 1734, S. 515/26.

[28]) STIRLING[26]), *prop.* 2, Beispiel 6 (S. 28/29); 9 richtige Dezimalen von s_2, ferner *prop.* 11, Beispiel 1 (S. 54/55; 17 richtige Dezimalen) und *prop.* 12, Beispiel 2 (S. 64/66; wiederum 17 richtige Dezimalen).

[29]) Das Original dieses Briefes ist verschollen, wir haben jedoch eine für die Akademie angefertigte Abschrift (Archiv der Moskauer Akademie), weil EULER in deren Auftrag an STIRLING schrieb und auch den Brief schon vor der Abfertigung am 4. [15.] 6. 1736 zur Genehmigung vorlegte.

und zwar mit genau dem nämlichen Zahlenfehler, den der Erstdruck von E 55 enthält. Dieser Fehler hätte die Wiedergabe von 20 richtigen Dezimalen verhindert! Also kann die Berechnung von s_2 in E 47 nicht vermittels der Summenformel vorgenommen worden sein.

Übrigens hat Euler seinen Irrtum alsbald bemerkt. Schon im zweiten Brief an Stirling vom 27. 7. [7. 8.] 1738 findet sich der richtige Wert vor[30]). Wir sehen übrigens hieraus, daß schon 1738 keine Möglichkeit mehr bestand, den bereits zum Teil ausgedruckten Band 8 der *Commentarii* textlich zu verändern, obwohl der Band erst 1741 ausgegeben wurde. Unklar bleibt freilich, warum keine Berichtigung am Ende des Bandes vorgenommen wurde.

8. Im Vorstehenden wurde versucht, das Hauptsächlichste über die Entstehungsgeschichte der Eulerschen Summenformel in gedrängter Form wiederzugeben. Noch ist manches ungeklärt, was wohl erst bei Herausgabe des gesamten Euler-Briefwechsels in helleres Licht gesetzt werden kann. Erst dann, wenn das noch unedierte Briefmaterial im Druck zugänglich und mit den nötigen Registern versehen ist [die z.B. beklagenswerter Weise in der *Correspondance*[6]) von Fuss fehlen], wird sich der Werdegang des Eulerschen Genius wirklich erkennen lassen — eine Angelegenheit, die mir von höchstem psychologischem Interesse zu sein scheint.

[30]) Das Original ist abgedruckt in Ch. Tweedie: *J. Stirling, a sketch of his life and works along with his scientific correspondence*, Oxford 1922, S. 181/91. Die fragliche Stelle befindet sich auf S. 183. Dort ist auch noch das Glied mit der 19. Ableitung richtig wiedergegeben. Im Druck ist der Fehler erstmals berichtigt in den *Institutiones calculi differentialis cum ejus usu in analysi finitorum ac doctrina serierum*, Berlin 1755 (E 212 von 1748) = *EO* X, ed. G. Kowalewski, Leipzig-Berlin 1913, S. 320.

(13b) Ichenhausen/Schw., Heinr. Sinzstr. 15

(Eingegangen am 29. September 1956)

LENSE, J.
Math. Zeitschr. Bd. 67, S. 147—152 (1957)

Bemerkungen zur Verwendung der Elementarteiler bei Paaren von nicht zerfallenden Kegelschnitten

Herrn GEORG FABER zum 80. Geburtstag am 5. April 1957 gewidmet

Von

JOSEF LENSE

Um die Theorie der Elementarteiler auf Kegelschnittspaare oder Kegelschnittsbüschel anwenden zu können, genügt es nicht, nur zu beweisen, daß zwei λ-Matrizen ersten Grades dann und nur dann äquivalent sind (d.h. auseinander durch linksseitige und rechtsseitige Multiplikation mit Matrizen hervorgehen, deren Elemente von λ nicht abhängen und deren Determinanten nicht verschwinden), wenn sie gleiche Elementarteiler besitzen; sondern man hat darüber hinaus noch zu zeigen, daß bei symmetrischen λ-Matrizen die beiden multiplizierenden Matrizen außerdem noch so gewählt werden können, daß die eine die transponierte der anderen ist. Beides erfordert immerhin einen nicht unbeträchtlichen Aufwand an Beweismitteln. So bedarf man z.B. zum Beweis des zweiten, sich auf symmetrische Matrizen beziehenden Teils der Behauptung der Tatsache, daß man, kurz gesagt, aus einer Matrix die Quadratwurzel ziehen kann, oder genauer ausgedrückt, daß sich zu einer Matrix, deren Determinante nicht verschwindet, eine ebensolche gleicher Ordnung angeben läßt, deren Quadrat die gegebene Matrix ist[1]).

Es dürfte daher von Vorteil sein, wenn es sich z.B. für Vorlesungszwecke nur um die projektive Einteilung der Kegelschnittspaare mit Hilfe ihrer Elementarteiler handelt, einen kürzeren Weg aufzufinden. Ein solcher soll im folgenden angegeben werden.

§ 1

Wir betrachten eine sog. λ-Matrix ersten Grades, d.h. eine Matrix von der Gestalt $\|a_{\mu\nu}+\lambda b_{\mu\nu}\|$ $(\mu,\nu=1,2,\ldots,n)$, wobei die $a_{\mu\nu}$ und $b_{\mu\nu}$ beliebige komplexe Zahlen sind und λ ein komplexer Parameter sein soll.

Ihre r-reihigen Unterdeterminanten sind Polynome in λ mit einem Grad $\leq r$. Wir wollen ihren größten gemeinsamen Teiler so normieren, daß der Koeffizient der höchsten Potenz von λ gleich 1 wird. Dann ändert sich bekanntlich[2]) dieser größte gemeinschaftliche Teiler nicht, wenn wir die λ-Matrix links oder rechts mit einer Matrix $\|c_{\mu\nu}\|$ multiplizieren, wobei auch die $c_{\mu\nu}$ komplexe Zahlen mit nicht verschwindender Determinante sein sollen.

[1]) Siehe z.B. BÔCHER, M.: Einführung in die höhere Algebra, S. 282—336. Deutsch von H. BECK. Leipzig u. Berlin: B. G. Teubner 1910.

[2]) Siehe z.B. M. BÔCHER, a. a. O., S. 297.

Es seien jetzt die $a_{\mu\nu}$ und $b_{\mu\nu}$ die Koeffizienten der Gleichungen zweier nicht zerfallenden Kegelschnitte in homogenen Dreieckskoordinaten der komplexen projektiven Ebene $\sum_{\mu,\nu=1}^{3} a_{\mu\nu}x_\mu x_\nu = 0$ und $\sum_{\mu,\nu=1}^{3} b_{\mu\nu}x_\mu x_\nu = 0$, die aus ihnen gebildeten Determinanten also von Null verschieden und $a_{\mu\nu} = a_{\nu\mu}$, $b_{\mu\nu} = b_{\nu\mu}$. Üben wir auf die x_μ eine lineare homogene Transformation mit nicht verschwindender Determinante aus, $x_\mu = \sum_{\nu=1}^{3} c_{\mu\nu}x'_\nu$ $(\mu = 1, 2, 3)$, wir können sie als Kollineation oder Koordinatentransformation deuten, so entsteht die Matrix der Koeffizienten in den Koordinaten x'_μ aus der alten, in dem man rechts mit der Matrix $\|c_{\mu\nu}\|$ und links mit ihrer transponierten multipliziert. Dabei ändert sich, wie eben erwähnt, der entsprechend normierte größte gemeinsame Teiler der r-reihigen Unterdeterminanten der λ-Matrix $\|a_{\mu\nu} + \lambda b_{\mu\nu}\|$ nicht.

§ 2

Wir wollen nun die einzelnen Fälle durchgehen.

1. Die Elemente $a_{\mu\nu} + \lambda b_{\mu\nu}$ haben einen gemeinsamen Teiler. Weil sie vom Grad 1 in λ sind, hat er die Gestalt $\lambda - \alpha$. Die Unterdeterminanten haben daher den größten gemeinsamen Teiler $(\lambda - \alpha)^2$, die Determinante selbst ist bis auf einen von λ unabhängigen Faktor $(\lambda - \alpha)^3$.

2. Die Elemente $a_{\mu\nu} + \lambda b_{\mu\nu}$ haben keinen gemeinsamen Teiler, wohl aber die Unterdeterminanten einen solchen ersten Grades $\lambda - \alpha$. Dieser ist nach dem Entwicklungsatz der Determinanten auch Teiler der Determinante, sie ist also bis auf einen von λ unabhängigen Faktor von der Gestalt $(\lambda - \alpha)(\lambda - \beta)(\lambda - \gamma)$. Weil aber die Determinante der Unterdeterminanten einerseits durch $(\lambda - \alpha)^3$ teilbar, andererseits dem Quadrat der ursprünglichen Determinante gleich, somit bis auf einen von λ unabhängigen Faktor $(\lambda - \alpha)^2(\lambda - \beta)^2(\lambda - \gamma)^2$ ist, muß entweder $\alpha = \beta \neq \gamma$ oder $\alpha = \beta = \gamma$ sein.

3. Die Elemente haben keinen gemeinsamen Teiler, die Unterdeterminanten einen solchen zweiten Grades $(\lambda - \alpha)(\lambda - \beta)$. Dieser Fall kann nicht eintreten. Denn dann wäre die Determinante der Unterdeterminanten bis auf einen von λ unabhängigen Faktor $(\lambda - \alpha)^3(\lambda - \beta)^3$. Da aber diese Determinante das Quadrat der ursprünglichen, also bis auf einen von λ unabhängigen Faktor $(\lambda - \alpha)^2(\lambda - \beta)^2(\lambda - \gamma)^2$ ist, müßte $\gamma = \beta = \alpha$ sein. Bildet man nun die Unterdeterminanten der Determinante der Unterdeterminanten, so wären diese einerseits bis auf einen von λ unabhängigen Faktor von der Gestalt $(\lambda - \alpha)^4$, andererseits nach einem bekannten Determinantensatz[3]) bis auf einen von λ unabhängigen Faktor von der Gestalt $(a_{\mu\nu} + \lambda b_{\mu\nu})(\lambda - \alpha)^3$, was unmöglich ist, da die Elemente $a_{\mu\nu} + \lambda b_{\mu\nu}$ nach Voraussetzung keinen gemeinsamen Teiler haben.

4. Die Elemente und die Unterdeterminanten haben keinen gemeinsamen Teiler, die Determinante ist bis auf einen von λ unabhängigen Faktor von der

[3]) Siehe z. B. M. BÔCHER, a. a. O., S. 33.

Gestalt $(\lambda-\alpha)(\lambda-\beta)(\lambda-\gamma)$, wobei insbesondere auch $\alpha=\beta\neq\gamma$ oder $\alpha=\beta=\gamma$ sein kann.

Man erkennt demnach, daß in allen möglichen Fällen immer der größte gemeinsame Teiler der Unterdeterminanten durch den größten gemeinsamen Teiler der Elemente und die Determinante durch den größten Teiler der Unterdeterminanten teilbar ist. Die hierbei auftretenden Quotienten und den größten gemeinsamen Teiler der Elemente bezeichnet man, wenn sie von λ wirklich abhängen, als invariante Faktoren der λ-Matrix, die Potenzen der verschiedenen Linearfaktoren in den invarianten Faktoren als Elementarteiler.

Nach der eben durchgeführten Fallunterscheidung sind demnach folgende sechs Fälle möglich:

1. Elemente und Unterdeterminanten haben keinen gemeinsamen Teiler, die Determinante ist von der Gestalt $(\lambda-\alpha)(\lambda-\beta)(\lambda-\gamma)$ mit paarweise verschiedenen α, β, γ, die Elementarteiler sind $\lambda-\alpha, \lambda-\beta, \lambda-\gamma$.

2. Wie bei 1., aber mit $\alpha=\beta$, die Elementarteiler sind $(\lambda-\alpha)^2, (\lambda-\gamma)$.

3. Wie bei 1., aber mit $\alpha=\beta=\gamma$, ein einziger Elementarteiler $(\lambda-\alpha)^3$.

4. Die Elemente haben keinen gemeinsamen Teiler, die Unterdeterminanten den größten gemeinsamen Teiler $\lambda-\alpha$, die Determinante ist von der Gestalt $(\lambda-\alpha)^2(\lambda-\beta)$ mit $\beta\neq\alpha$ die Elementarteiler sind $\lambda-\alpha, \lambda-\alpha, \lambda-\beta$.

5. Wie bei 4., aber mit $\beta=\alpha$, die Elementarteiler sind $\lambda-\alpha, (\lambda-\alpha)^2$.

6. Die Elemente haben den größten gemeinsamen Teiler $\lambda-\alpha$, die Unterdeterminanten $(\lambda-\alpha)^2$, die Determinante ist von der Gestalt $(\lambda-\alpha)^3$, die Elementarteiler sind $\lambda-\alpha, \lambda-\alpha, \lambda-\alpha$.

§ 3

Nach den vorhergegangenen Erörterungen ändern sich die invarianten Faktoren und Elementarteiler nicht, wenn man die λ-Matrix mit einer Matrix $\|c_{\mu\nu}\|$ multipliziert, deren Determinante von Null verschieden ist, also z. B. wenn man auf die Koordinaten x_μ eine homogene lineare Transformation mit nicht verschwindender Determinante ausübt, d. h. die Kegelschnittspaare einer Kollineation oder Koordinatentransformation unterwirft. Da es aber dabei nicht auf die quadratischen Formen, sondern nur auf die Gleichungen des Kegelschnittpaares ankommt, also die Koordinaten x_μ in den beiden Kegelschnittsgleichungen noch je mit einem bei beiden Gleichungen verschiedenen Proportionalitätsfaktor ϱ_1 bzw. ϱ_2 multipliziert werden können, hat die Matrix nach der Transformation die Gestalt $\left\| \sum_{\sigma,\tau=0}^{3} c_{\sigma\mu}(\varrho_1^2 a_{\sigma\tau}+\lambda\varrho_2^2 b_{\sigma\tau})c_{\tau\nu}\right\|$, d. h. die Konstanten α, β, γ der Elementarteiler werden alle gleichzeitig mit einem gemeinsamen Faktor $\left(\frac{\varrho_1}{\varrho_2}\right)^2$ multipliziert. Projektive Invarianten des Kegelschnittpaares sind hiernach die Exponenten der Elementarteiler und die Gleichheit oder Verschiedenheit der Größen α, β, γ, die Größen selbst nur bis auf einen gemeinsamen Faktor.

Man hat daher folgende sechs Fälle zu unterscheiden:

1. Drei paarweise verschiedene Elementarteiler, Bezeichnung (die sog. Charakteristik) [1 1 1].

2. Das Quadrat eines linearen Elementarteilers und ein davon verschiedener linearer Elementarteiler, [2 1].

3. Der Kubus eines linearen Elementarteilers, [3].

4. Zwei gleiche lineare und ein davon verschiedener linearer Elementarteiler, [(1 1) 1].

5. Ein linearer Elementarteiler und sein Quadrat, [(2 1)].

6. Drei gleiche lineare Elementarteiler, [(1 1 1)].

Ihnen entsprechen sechs Klassen von Kegelschnittspaaren in der Weise, daß bei einer projektiven Transformation die Kegelschnittspaare einer Klasse immer in solche derselben Klasse übergehen.

Diese sechs Klassen lassen sich aber auch geometrisch durch die Schnittpunkte der Kegelschnitte eines Paares kennzeichnen. Um zu den Schnittpunkten, also analytisch gesprochen, zu den Lösungen der beiden homogenen Gleichungen zweiten Grades zu gelangen, durch welche die beiden Kegelschnitte gegeben sind, kann man, ohne den Begriff der Resultante oder das Theorem von BÉZOUT heranziehen zu müssen, in folgender Weise vorgehen: Durch passende Wahl des Koordinatendreiecks und des Einheitspunktes läßt sich erreichen, daß die Gleichung des einen Kegelschnittes die Form $x_1 x_2 = x_3^2$ annimmt. Berechnet man aus dieser Gleichung $x_1 = \frac{x_3^2}{x_2}$ und setzt diesen Wert in die Gleichung des zweiten Kegelschnitts ein, so ergeben sich eine Gleichung vierten Grades für das Verhältnis $x_2 : x_3$ und damit, je nach den Lösungsmöglichkeiten dieser Gleichung, folgende sechs Fälle für die gegenseitige Lage der beiden Kegelschnitte:

1. Vier getrennte Schnittpunkte.

2. Zwei der vier Schnittpunkte fallen zusammen.

3. Drei der vier Schnittpunkte fallen zusammen.

4. Je zwei der vier Schnittpunkte fallen zusammen.

5. Alle vier Schnittpunkte fallen zusammen.

6. Die beiden Kegelschnitte haben sämtliche Punkte gemeinsam.

Bei projektiver Transformation bleiben die sechs Klassen getrennt, d.h. Kegelschnittspaare einer solchen Klasse gehen wieder in Paare derselben Klasse über. Sie entsprechen den obigen sechs Fällen der Elementarteiler. Wir können nämlich in jeder der sechs Schnittpunktsklassen die Gleichungen irgendeines Kegelschnittspaares durch eine passende Koordinatentransformation auf eine geeignete Normalform bringen, wobei sich nach dem Obigen die Charakteristik nicht ändert. Für jede der sechs Normalformen stellt sich genau eine der sechs Charakteristiken heraus.

§ 4

1. Wir wählen das zu dem vollständigen Viereck der vier Schnittpunkte gehörige Diagonaldreieck als Koordinatendreieck. Bei passender Wahl des Einheitspunktes ergibt sich das Gleichungspaar $x_1^2+x_2^2=x_3^2$, $a\,x_1^2+b\,x_2^2=x_3^2$, wobei die Konstanten a und b von 0 und 1 und untereinander verschieden sind. Sie lassen sich nicht weiter normieren, sind also projektive Invarianten. Man erhält nämlich für das Doppelverhältnis d_1 der vier Schnittpunkte der Geraden $x_1=0$ mit den beiden Kurven $d_1=\left(\frac{1+\sqrt{b}}{1-\sqrt{b}}\right)^2$, also $b=\left(\frac{1-\sqrt{d_1}}{1+\sqrt{d_1}}\right)^2$, ebenso für die vier Schnittpunkte der Geraden $x_2=0$ mit den Kegelschnitten $d_2=\left(\frac{1+\sqrt{a}}{1-\sqrt{a}}\right)^2$, $a=\left(\frac{1-\sqrt{d_2}}{1+\sqrt{d_2}}\right)^2$, bei $x_3=0$ schließlich $d_3=\left(\frac{\sqrt{a}+\sqrt{b}}{\sqrt{a}-\sqrt{b}}\right)^2=\left(\frac{1-\sqrt{d_1 d_2}}{\sqrt{d_1}-\sqrt{d_2}}\right)^2$ oder $1+\sqrt{d_2 d_3}+\sqrt{d_3 d_1}+\sqrt{d_1 d_2}=0$ bei entsprechender Reihenfolge der Schnittpunkte und passender Wahl des Zeichens der Quadratwurzeln. Die Elementarteiler sind $\lambda+1$, $\lambda+a$, $\lambda+b$, also Charakteristik [1 1 1].

2. Wir legen den Punkt $(1, 0, 0)$ in den Punkt A, in dem die beiden Schnittpunkte zusammenfallen und die Seite $x_2=0$ in die gemeinsame Tangente in A. B und C seien die beiden übrigen Schnittpunkte. In den Schnittpunkt der Tangente mit der Geraden BC legen wir den Punkt $(0, 0, 1)$. Dann ist der zur Tangente durch die beiden Geraden AB und AC getrennte vierte harmonische Strahl die Polare des Punktes $(0, 0, 1)$ in bezug auf die beiden Kegelschnitte. Ziehen wir also vom Punkt $(0, 0, 1)$ die zweite Tangente an einen der beiden Kegelschnitte, so liegt der Berührungspunkt auf der genannten Polaren. In diesen Punkt legen wir den dritten Eckpunkt $(0, 1, 0)$ des Koordinatendreiecks. Dann ergeben sich bei passender Wahl des Einheitspunktes für die beiden Kurven die Gleichungen $x_1 x_2=x_3^2$, $a x_1 x_2+x_2^2=x_3^2$ mit $a\neq 0, 1$. Die Geraden AB und AC treffen die Seite $x_1=0$ des Koordinatendreiecks in den Punkten D und E. Wir verbinden diese Punkte mit dem Schnittpunkt F der Geraden BC und $x_3=0$. Die Geraden DF und EF schneiden die beiden Kegelschnitte in vier Punkten gleichen Doppelverhältnisses d. Man erhält nämlich für die Punkte B und C die Koordinaten $\left(1, 1-a, \pm\sqrt{1-a}\right)$, für die Geraden AB, AC und BC die Gleichungen $x_2=\pm\sqrt{1-a}\,x_3$, $x_2=(1-a)\,x_1$, für die Punkte D, E, F die Koordinaten $\left(0, \pm\sqrt{1-a}, 1\right)$, $(1, 1-a, 0)$, schließlich $d=-\left[\frac{a+4-\sqrt{5(a^2+4)}}{2(a-1)}\right]^2$, $a=\frac{1+d-4\sqrt{-d}}{1+d+\sqrt{-d}}$ ist projektive Invariante. Die Elementarteiler sind $\lambda+1$, $(\lambda+a)^2$, also Charakteristik [2 1].

3. Wir legen den Punkt $(1, 0, 0)$ in den Punkt A, in dem die drei Schnittpunkte zusammenfallen, die Seite $x_2=0$ in die gemeinsame Tangente in A und den Punkt $(0, 1, 0)$ in den vierten Schnittpunkt B. Die Seite $x_1=0$ sei die Tangente in B an einen der beiden Kegelschnitte. Damit erhält man

bei passender Wahl des Einheitspunktes das Gleichungspaar $x_1x_2=x_3^2$, $x_1x_2+x_2x_3=x_3^2$, $(\lambda+1)^3$ ist Elementarteiler, Charakteristik [3].

4. Wir legen die Punkte (1, 0, 0) und (0, 1, 0) in die beiden Berührungspunkte, die Seiten $x_1=0$ und $x_2=0$ in die entsprechenden gemeinsamen Tangenten. Dann ergibt sich bei passender Wahl des Einheitspunktes das Gleichungspaar $x_1x_2=x_3^2$, $ax_1x_2=x_3^2$ mit $a\neq 0, 1$. Jede Gerade durch den Punkt (0, 0, 1) schneidet die beiden Kegelschnitte in Punkten gleichen Doppelverhältnisses $d=\left(\frac{1+\sqrt{a}}{1-\sqrt{a}}\right)^2$, $a=\left(\frac{1-\sqrt{d}}{1+\sqrt{d}}\right)^2$ ist projektive Invariante, die Elementarteiler sind $\lambda+1$, $\lambda+a$, $\lambda+a$, die Charakteristik ist [(1 1) 1].

5. Wir legen den Punkt (1, 0, 0) in den gemeinsamen Berührungspunkt, die Seite $x_2=0$ in die gemeinsame Tangente, die Seite $x_1=0$ in eine andere Tangente eines der beiden Kegelschnitte und den Punkt (0, 1, 0) in den Berührungspunkt dieser Tangente. Die Gerade $x_1=0$ wird von den Kegelschnitten des durch die beiden gegebenen Kegelschnitte bestimmten Büschels in den Punkten einer Involution geschnitten, deren Doppelpunkte die Punkte (0, 1, 0) und (0, 0, 1) sind, weil die Doppelgerade $x_2^2=0$ dem Büschel angehört. Diese beiden Punkte sind also in bezug auf den zweiten Kegelschnitt konjugiert. Daher ergeben sich für die beiden Kegelschnitte bei passender Wahl des Einheitspunktes die Gleichungen $x_1x_2+x_3^2=0$, $x_1x_2+x_2^2=x_3^2$. Die Elementarteiler sind $\lambda+1$, $(\lambda+1)^2$, die Charakteristik ist [(2 1)].

6. Wir legen das Koordinatendreieck in ein Polardreieck der beiden zusammenfallenden Kegelschnitte und erhalten bei passender Wahl des Einheitspunktes für beide die Gleichung $x_1^2+x_2^2=x_3^2$, die Elementarteiler $\lambda+1$, $\lambda+1$, $\lambda+1$, somit die Charakteristik [(1 1 1)].

Die Einteilung der Kegelschnittspaare in sechs Klassen in bezug auf ihre Elementarteiler entspricht also der Reihenfolge nach genau ihrer Einteilung in sechs Klassen in bezug auf ihre Schnittpunkte. Ist demnach die Charakteristik des Paares bekannt, so kann man durch eine passende Kollineation oder Koordinatentransformation immer eine der obigen Normalformen[4]) herstellen.

Die beiden quadratischen Gleichungen der Kegelschnitte haben also z. B. dann und nur dann dieselben Lösungen, wenn der Fall 6. vorliegt, d. h. die Elemente der λ-Matrix den größten gemeinsamen Teiler $\lambda-\alpha$ haben, die Koeffizienten der beiden Gleichungen somit proportional sind.

[4]) Sie finden sich schon in dem Buch von T. J. I'A. BROMWICH, Quadratic forms and their Classification by means of Invariant-factors, Cambridge Tracts in Mathematics and Mathematical Physics Nr. 3, S. 32—35. Cambridge: University Press 1906.

München, Mathematisches Institut der Technischen Hochschule

(Eingegangen am 10. März 1956)

Lenz, H.
Math. Zeitschr. Bd. 67, S. 153—175 (1957)

Über die elliptischen Funktionen von Jacobi

Herrn Georg Faber zum 80. Geburtstag am 5. April 1957 gewidmet

Von

Hanfried Lenz

Einleitung

Herr Faber hat 1920 im Vorwort der von ihm völlig neu bearbeiteten Burkhardtschen Vorlesungen über elliptische Funktionen[1]) geschrieben: „Als ich nach dem allzufrühen Tode Burkhardts auf Wunsch des Verlags die Besorgung der dritten Auflage dieses Buchs übernahm, erblickte ich meine Aufgabe vor allem darin, die der Burkhardtschen Darstellung eigentümlichen Vorzüge zu erhalten, nämlich die enge Verknüpfung und Durchdringung Abelscher und Jacobischer, Cauchyscher, Riemannscher und Weierstrassscher Gedanken, sowie die bewußte Heranziehung der geometrischen Anschauung[2]) bei aller sonstigen Strenge im Analytischen."

Diesem Ziel will auch die vorliegende Arbeit dienen. Wer allerdings über eine so alte und nach allgemeiner Ansicht abgeschlossene Theorie überhaupt noch eine Arbeit schreibt, muß das begründen: Ein Vergleich der elliptischen Funktionen von Jacobi und Weierstrass zeigt, daß die Weierstraßschen Funktionen einen besonders einfachen systematischen Aufbau der Theorie ermöglichen, während sich die Jacobischen bis heute in den Anwendungen, insbesondere in den Funktionentafeln, behauptet haben[3]). E. H. Neville[4]) hat daher einen sehr bemerkenswerten Versuch gemacht, die Vorzüge beider Theorien zu vereinigen, d.h. eine Theorie der Jacobischen *Funktionen* auf Grund Weierstrassscher *Gedanken* zu geben und insbesondere den Jacobischen analoge Funktionen für beliebig gegebenes Periodengitter einzuführen. Die vorliegende Arbeit will den Nevilleschen Aufbau der Theorie noch vereinfachen; insbesondere soll gezeigt werden, daß es nicht nötig ist, die Weierstraßschen Funktionen und viele neue Bezeichnungen einzuführen, wie Neville tut, sondern daß bei unwesentlich modifizierter Auffassung die alten, von Gudermann und Glaisher eingeführten Bezeichnungen für die Jacobischen

[1]) Burkhardt-Faber: Elliptische Funktionen, 3. Aufl., Berlin u. Leipzig 1920.

[2]) Anschaulich-geometrische Betrachtungen zur Lösung des Umkehrproblems mit besonders einfachen Mitteln finden sich auch bei R. Fricke, Enc. math. Wiss. II B 3, insbes. S. 254ff., ferner: Die elliptischen Funktionen und ihre Anwendungen I, S. 164ff., Leipzig u. Berlin 1916. Die im folgenden durchgeführte Lösung des Umkehrproblems schließt sich an Frickesche Gedankengänge an.

[3]) Vgl. etwa Milne-Thomson: Jacobian Elliptic Function Tables. New York 1950. — Byrd, P. F., and Friedman, M. D.: Handbook of Elliptic Integrals for Engineers and Physicists. Berlin-Göttingen-Heidelberg 1954. Schuler, M., u. H. Gebelein: Fünfstellige Tabellen zu den elliptischen Funktionen. — Acht- und neunstellige Tabellen zu den elliptischen Funktionen. Springer 1955.

[4]) Neville, E. H.: Jacobian Elliptic Functions, 2nd ed. Oxford 1951.

Funktionen, ihre Reziproken und Quotienten völlig ausreichen. Die Thetafunktionen normiert man nach dem Muster NEVILLEs besser etwas anders als üblich. Man erhält so ein ganz einfaches Verhalten gegenüber Periodentransformationen. Die Jacobischen elliptischen Funktionen in der hier behandelten Auffassung sind ein Sonderfall der meromorphen Funktionen mit doppeltperiodischem *absolutem Betrag* und genau einem Pol im Periodenparallelogramm (des Betrages). Diese letzteren Funktionen geben einen kaum noch zu vereinfachenden Zugang zur Transformation höheren Grades der elliptischen Funktionen.

Die Weierstraßschen Funktionen treten in der hier gewählten Darstellung kaum mehr in Erscheinung. Man darf aber deshalb nicht denken, sie seien in der Theorie der elliptischen Funktionen ganz entbehrlich. Das Verhalten der Jacobischen Funktionen hinsichtlich Periodentransformationen wird bei den Jacobischen Funktionen und Thetafunktionen in der hier vertretenen Auffassung zwar sehr einfach; völlige Invarianz läßt sich aber nur bei den Weierstraßschen Funktionen erreichen.

§ 1. Elliptische Integrale erster Gattung

1. Konforme Abbildung. Das elliptische Integral 1. Gattung

$$u = \text{const} \cdot \int \frac{dz}{\sqrt{\prod_{\nu}(z - z_\nu)}} \qquad (\nu = 1, 2, 3, 4 \text{ oder } \nu = 1, 2, 3) \tag{1.1}$$

wird bekanntlich durch jede ganze oder gebrochene lineare Transformation der Variablen z in ein ebensolches übergeführt. Es ist daher keine Beschränkung der Allgemeinheit, es in der sog. Riemannschen Normalform

$$\mu u = \int \frac{dZ}{2\sqrt{Z(1-Z)(1-\lambda Z)}} = w \tag{1.2}$$

zu schreiben. Dabei sei

$$K = \int_0^1, \qquad -iK' = \int_{-\infty}^0 \tag{1.3}$$

gesetzt, wobei geradlinig zu integrieren ist und das Vorzeichen der Wurzel im Integranden so festgelegt sein soll, daß der Integrand gegen $+\infty$ strebt, wenn Z von rechts her auf der reellen Achse gegen 0 strebt, und gegen $-i\infty$, wenn Z von links gegen 0 strebt. Wenn $1/\lambda$ nicht <1 ist, was sich bei der Herstellung der Normalform (1.2) erreichen läßt, sind die „vollständigen Integrale" (1.3) damit eindeutig bestimmt. $2K$ und $2iK'$ sind nach dem Cauchyschen Satz gleich den Integralen auf geschlossenen Wegen, die die Strecken $[0, 1]$ bzw. $[1, 1/\lambda]$ und keine weiteren Verzweigungspunkte des Integranden umlaufen, wobei die Durchlaufungsrichtung noch von der Wahl des Vorzeichens der Wurzel abhängt.

Durch je drei der vier Verzweigungspunkte des Integranden von (1.1) geht ein Kreis, der die Zahlenkugel in zwei Kalotten zerlegt, von denen

mindestens eine den vierten Verzweigungspunkt nicht im Innern enthält. Es gilt [Fig. 1]

Satz 1. *$\mathfrak{K}$ sei eine Kalotte der Zahlenkugel, die auf dem Rande drei Verzweigungspunkte des Integrals* (1.1) *enthält, während im Innern kein solcher Verzweigungspunkt liegt. Dann bildet das Integral* (1.1) *die Kalotte $\mathfrak{K}$ auf ein konvexes Kurvendreieck $\mathfrak{D}$ mit drei rechten Winkeln ab, dessen Ecken sich zu einem geradlinigen Dreieck verbinden lassen, das nur rechtwinklig ist, wenn alle vier Verzweigungspunkte auf einem Kreis liegen und sonst spitzwinklig. In dem Grenzfall, daß alle vier Verzweigungspunkte auf einem Kreis liegen, geht das Kurvendreieck in ein Rechteck über* [Fig. 2].

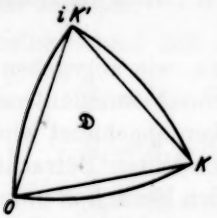


Fig. 1

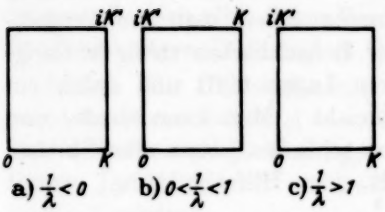


Fig. 2 a—c

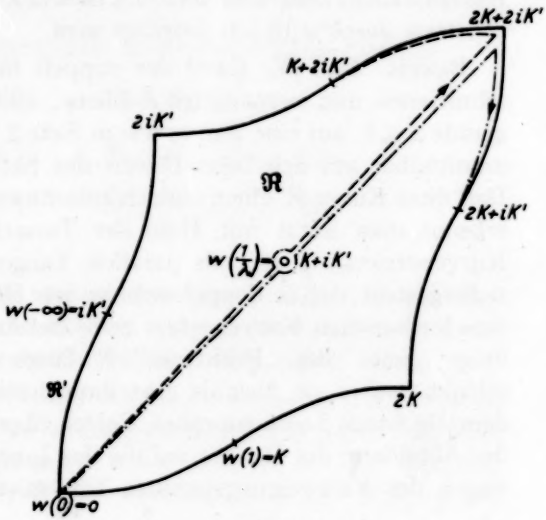


Fig. 3

Beweis: Es genügt, die Normalform (1.2) zu betrachten, in der wir noch $\mathfrak{Im}\,\lambda \geq 0$ annehmen dürfen (andernfalls führe man $1/Z$ als neue Variable ein). $\mathfrak{K}$ wird dann die obere Halbebene. $\arg \frac{1}{\sqrt{1-\lambda Z}}$ nimmt im Fall $\mathfrak{Im}\,\lambda > 0$ stetig und monoton um insgesamt 90° zu, wenn Z die reelle Achse (als Rand der oberen Halbebene) von $-\infty$ bis $+\infty$ durchläuft, während $\arg \frac{1}{\sqrt{Z}}$ und $\arg \frac{1}{\sqrt{1-Z}}$ stückweise konstant sind und nur bei $z=0$ bzw. $z=1$ um $+90°$ springen. Das Bild der reellen Achse wird also eine aus drei stetig differenzierbaren konvexen Bögen zusammengesetzte Kurve mit drei rechten Eckenwinkeln und einer Gesamtbiegung der Bögen von 90°. Daher müssen die Winkel des von den drei Ecken gebildeten geradlinigen Dreiecks spitz sein. Keine Tangente eines der drei Bögen ist parallel zu einer Tangente eines der drei anderen Bögen. Die Ecken liegen bei iK', 0, K.

Im Fall $\mathfrak{Im}\,\lambda = 0$ werden aus zweien der Bögen Strecken, aus dem dritten ein rechtwinkliges Streckenpaar. Weil u im Innern von $\mathfrak{K}$ regulär ist, können

wir in bekannter Weise[5]) von der Abbildung des Randes auf die Abbildung des Inneren schließen, womit Satz 1 bewiesen ist. Ebenso folgt [Fig. 3]

Satz 2. *Es sei entweder* $\Im\mathrm{m}\ \lambda > 0$ *oder* $\frac{1}{\lambda} > 1$. *Dann bildet das Integral* (1.2) *die doppelt überdeckte, längs der Halbgeraden* $Z \leq 1$ *aufgeschnittene und bei* $Z = 1/\lambda$ *verzweigte Z-Ebene umkehrbar eindeutig und (außer an der Stelle* $1/\lambda$*) konform auf ein Kurvenrechteck* $\mathfrak{R}$ *ab, das von vier einzeln zentralsymmetrischen und paarweise kongruenten und homothetischen Kurvenbögen begrenzt wird. Jeder dieser Kurvenbögen besteht aus zwei kongruenten und gegeneinander um 180° verdrehten Teilbögen, die je einem der Bögen des in Satz 1 beschriebenen Kurvendreiecks kongruent sind. Es ist* $w(1/\lambda) = K + iK'$, *wenn die Integrationskonstante durch* $w(0) = 0$ *festgelegt wird.*

Beweis. Daß der Rand der doppelt überdeckten, wie angegeben aufgeschnittenen und verzweigten Z-Ebene, also die viermal durchlaufene Halbgerade $Z \leq 1$, auf eine Kurve wie in Satz 2 beschrieben abgebildet wird, folgt unmittelbar aus den beim Beweis des Satzes 1 angestellten Betrachtungen. Daß diese Kurve $\mathfrak{R}'$ einen einfach zusammenhängenden Bereich $\mathfrak{R}$ umschließt, erkennt man leicht mit Hilfe der Tatsache, daß verschiedene Bögen des Kurvendreiecks $\mathfrak{D}$ niemals parallele Tangenten haben. Denn dadurch wird sichergestellt, daß $\mathfrak{R}'$ doppelpunktfrei ist. [Man könnte auch ohne Verwendung des Jordanschen Kurvensatzes zum Ziel kommen, indem man eine krummlinige „Seite" des „Rechtecks" $\mathfrak{R}'$ längs der benachbarten stetig entlangschiebt, wobei sie niemals eine ihrer früheren Lagen trifft und daher ein dem Rechteck homöomorphes Gebiet überstreicht.] Man kann wieder von der Abbildung des Randes auf die des Inneren schließen [wenn man Skrupel wegen des Verzweigungspunktes hat, mittels einer Hilfsabbildung], womit Satz 2 bewiesen ist. Wegen $\int_{-\infty}^{0} = \pm \int_{1}^{1/\lambda}$ ist $w(1/\lambda) = K + iK'$.

2. Die Umkehrfunktion. Im Innern und auf dem Rand von $\mathfrak{R}$ ist daher Z als Funktion von w meromorph. Für Rand- und Eckpunkte folgt das aus der Reihenentwicklung für $w(Z)$ nach dem funktionentheoretischen Satz über Umkehrfunktionen. $Z(w)$ ist also in einem abgeschlossenen Bereich $\hat{\mathfrak{R}}$ meromorph, der $\mathfrak{R}$ in seinem Inneren enthält. Mit $\hat{\mathfrak{R}} + a$ sei die Menge aller komplexen Zahlen $r + a$ bezeichnet, für die r in $\hat{\mathfrak{R}}$ liegt. Nach Satz 2 und dem Eindeutigkeitssatz für analytische Funktionen gelten die Funktionalgleichungen

$$Z(w \pm 2K) = Z(w), \tag{1.4}$$

$$Z(w \pm 2iK') = Z(w), \tag{1.5}$$

soweit sie einen Sinn haben, d.h. soweit die Zahlen w und $w \pm 2K$ bzw. w und $w \pm 2iK'$ beide in $\hat{\mathfrak{R}}$ liegen. Ebenso ist $Z(2K + 2iK' - w) = Z(w)$, also $Z(w) = Z(-w)$.

[5]) Vgl. z.B. BIEBERBACH, L.: Einführung in die konforme Abbildung, S. 47. Sammlung Göschen Bd. 768. — OSGOOD: Lehrbuch der Funktionentheorie, Bd. I, 2. Aufl., S. 377. Leipzig u. Berlin 1912.

Diese Funktionalgleichungen können nun aber dazu verwendet werden, den Definitionsbereich der Funktion $Z(w)$ auf den Bereich $\hat{\mathfrak{R}}+2K$ so auszudehnen, daß $Z(w)$ auch in diesem Bereich bis auf Pole regulär und eindeutig wird. Die Bereiche $\hat{\mathfrak{R}}$ und $\hat{\mathfrak{R}}+2K$ haben einen einfach zusammenhängenden Bereich mit inneren Punkten gemein und daher ist $Z(w)$ auch in ihrem Vereinigungsbereich

$$\hat{\mathfrak{R}}_1 = \hat{\mathfrak{R}} \cup (\hat{\mathfrak{R}} + 2K)$$

bis auf Pole regulär und eindeutig. Denn jeder geschlossene Weg im Vereinigungsbereich läßt sich in endlich viele geschlossene Wege in den Einzelbereichen zerlegen.

Derselbe Schluß läßt sich nacheinander anwenden auf die Bereiche

$$\hat{\mathfrak{R}}_2 = \hat{\mathfrak{R}}_1 \cup (\hat{\mathfrak{R}}_1 + 2iK'), \qquad \hat{\mathfrak{R}}_3 = \hat{\mathfrak{R}}_2 \cup (\hat{\mathfrak{R}}_2 - 4K),$$

$$\hat{\mathfrak{R}}_4 = \hat{\mathfrak{R}}_3 \cup (\hat{\mathfrak{R}}_3 - 4iK'), \qquad \hat{\mathfrak{R}}_5 = \hat{\mathfrak{R}}_4 \cup (\hat{\mathfrak{R}}_4 + 8K),$$

$$\hat{\mathfrak{R}}_6 = \hat{\mathfrak{R}}_5 \cup (\hat{\mathfrak{R}}_5 + 8iK'), \qquad \hat{\mathfrak{R}}_7 = \hat{\mathfrak{R}}_6 \cup (\hat{\mathfrak{R}}_6 - 16K),$$

. . . . . .

$Z(w)$ läßt sich daher als eindeutige und meromorphe Funktion in den Vereinigungsbereich aller dieser Bereiche, *d.h. die volle w-Ebene*, fortsetzen, wobei die Beziehungen (1.4) und (1.5) erhalten bleiben. Wir haben damit

Satz 3. *Die Umkehrfunktion $Z(w)$ des elliptischen Integrals* (1.2) *ist in der ganzen w-Ebene eindeutig, meromorph und doppeltperiodisch. Ihre Nullstellen, Pole, 1-Stellen, $1/\lambda$-Stellen liegen genau bei $w \equiv 0,\ iK',\ K,\ K+iK'$* (mod $2K$, $2iK'$). *Daher sind $2K$ und $2iK'$ primitive Perioden von $Z(w)$. $Z(w)$ ist eine gerade Funktion.*

Die w-Ebene ist also schachbrettartig von den Bereichen $\hat{\mathfrak{R}} + 2mK + 2niK'$ überdeckt und jeder dieser Bereiche ist konformes Bild der, wie in Satz 2 beschrieben, aufgeschnittenen und verzweigten doppelt überdeckten Z-Ebene.

Man hätte die Eindeutigkeit von $Z(w)$ natürlich auch mit Hilfe des sog. Rationalitätsradius nach BIEBERBACH beweisen können[6]), doch wurde der vorliegende Weg gewählt, weil er besonders anschaulich ist.

§ 2. Das Periodenverhältnis

1. Die konforme Abbildung durch das Periodenverhältnis. Wir setzen

$$\tau = \frac{iK'}{K}. \tag{2.1}$$

Für $\lambda \neq 0, 1, \infty$ sind K und K' als bestimmte Integrale mit von λ unabhängigen Grenzen nach λ differenzierbar und nach Satz 1 von Null verschieden. Daher ist τ als Funktion von λ für $\lambda \neq 0, 1, \infty$ regulär.

[6]) BIEBERBACH, L.: Lehrbuch der Funktionentheorie, Bd. I, S. 239. Leipzig u. Berlin 1921. Vgl. auch SIEGEL, C. L.: Ausgewählte Fragen der Funktionentheorie. Göttinger Vorlesung 1953/54, S. 32ff.

Für $\Im\mathfrak{m}\,\lambda>0$ ist das Dreieck $0, K, iK'$ nach Satz 1 *spitzwinklig* und daher liegt τ in dem Spitzendreieck

$$(2.2)\qquad \left\{\begin{array}{l} 0<\Re\mathfrak{e}\,\tau<1, \\ |\tau-0{,}5|>0{,}5 \end{array}\right\} \text{ für } \Im\mathfrak{m}\,\lambda>0 \text{ [Fig. 4a], außerdem } \Im\mathfrak{m}\,\tau>0.$$

In den Grenzfällen a) $1/\lambda<0$, b) $0<1/\lambda<1$, c) $1/\lambda>1$ geht das Kurvendreieck $\mathfrak{D}$ in die Rechtecke der Fig. 2 über und daraus folgt

im Fall a) $\qquad \Re\mathfrak{e}\,\tau=1,$

im Fall b) $\qquad |K|=|-K+2iK'| \quad$ oder $\quad 0{,}5=|\tau-0{,}5|,$

im Fall c) $\qquad \Re\mathfrak{e}\,\tau=0.$

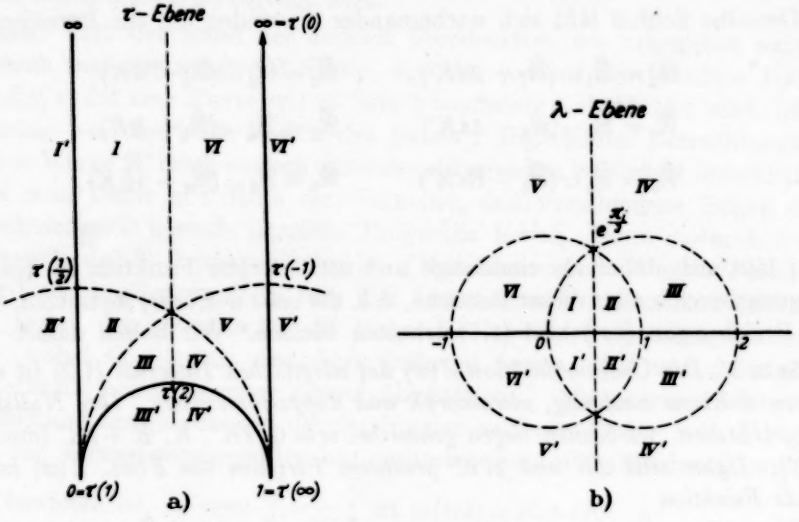


Fig. 4 a und b

Für *reelle* $\lambda\neq 0, 1, \infty$ ist ferner $\frac{d\tau}{d\lambda}\neq 0$, denn andernfalls würde eine gewisse Umgebung eines solchen Punktes λ so auf die τ-Ebene abgebildet, daß Punkte aus der oberen λ-Halbebene außerhalb des Bereiches (2.2) abgebildet würden, was nicht sein kann. Wenn also $1/\lambda$ die Geradenstücke $[-\infty, 0]$ bzw. $[0, 1]$ bzw. $[1, \infty]$ durchläuft, so durchläuft τ Teile der den Bereich (2.2) begrenzenden Kreisbögen bzw. Geradenstücke in konstanter Richtung. τ durchläuft diese Kreisbögen bzw. Geradenstücke sogar *ganz*. Um das zu zeigen, müssen wir noch beweisen, daß

$$(2.3)\qquad \lim_{\lambda\to 1}\tau(\lambda)=0, \quad \lim_{\lambda\to\infty}\tau(\lambda)=1, \quad \lim_{\lambda\to 0}\tau(\lambda)=\infty$$

ist.

Beweis. Bei geradliniger Integration längs des Randes der oberen Halbebene wird im Fall $\Im\mathfrak{m}\,\lambda>0$ nach (1.3) und Satz 1

$$(2.4)\qquad \int_{-\infty}^{0}=-iK', \quad \int_{0}^{1}=K, \quad \int_{1}^{\infty}=iK'-K.$$

Die Summe der drei Integrale verschwindet. Für $\lambda \to 1$ bleibt das erste Integral beschränkt. Wir zeigen, daß das zweite gegen ∞ strebt. Die Strecke $[0, 1]$ wird durch $u(Z)$ auf einen Kurvenbogen der Biegung $\leq 90°$ abgebildet, dessen Sehnenlänge mindestens das $1/\sqrt{2}$-fache seiner Bogenlänge beträgt. Daher ist

$$\sqrt{2}\,|K| \geq \int_0^1 \frac{dZ}{2\sqrt{Z(1-Z)\,|1-\lambda Z|}} \geq \frac{1}{2\sqrt{|\lambda|}} \int_0^1 \frac{dZ}{\sqrt{1\,(1-Z)\left(\left|\frac{1}{\lambda}-1\right|+1-Z\right)}}.$$

Mit λ strebt auch $c = \left|\frac{1}{\lambda} - 1\right| + 1$ gegen 1. Daher wird

$$2\sqrt{2|\lambda|}\,K \geq \int_0^1 \frac{dZ}{\sqrt{(1-Z)\,(c-Z)}} \geq \int_0^1 \frac{dZ}{\frac{1+c}{2}-Z}\,,$$

$$= \log\frac{1+c}{2} - \log\left(\frac{1+c}{2}-1\right) \to \infty \qquad \text{für} \qquad c \to 1.$$

Weil iK' beschränkt bleibt, folgt $\tau \to 0$. Die anderen beiden Gln. (2.3) ergeben sich genau so. Wie man z. B. durch lineare Transformation von Z nachprüfen kann, bleibt für $\lambda \to 0, 1, \infty$ immer dasjenige Integral (2.4), in dem der Grenzwert von $1/\lambda$ nicht als Grenze vorkommt, beschränkt, während die beiden anderen über alle Grenzen wachsen, und daraus folgt die Behauptung.

Durch $\tau(\lambda)$ wird also die reelle Achse als Rand der oberen Halbebene umkehrbar eindeutig und stetig auf den Rand des Bereiches (2.2) der Fig. 3 abgebildet, während $\tau(\lambda)$ in der oberen Halbebene regulär ist und innere Punkte auf innere Punkte von (2.2) abbildet. Daraus folgt

Satz 4. *Das Periodenverhältnis* $\tau(\lambda) = \frac{iK'(\lambda)}{K(\lambda)}$ *bildet die obere* λ*-Halbebene umkehrbar eindeutig und konform auf das Kreisbogendreieck* (2.2) *ab.*

Weil man in jedem Periodengitter bekanntlich zwei primitive Perioden finden kann, deren Verhältnis in der abgeschlossenen Hülle des Bereiches (2.2) [aber in keiner Ecke] liegt, folgt aus Satz 4, *daß zu jedem vorgegebenen Periodengitter eine nichtkonstante elliptische Funktion existiert, die Umkehrfunktion eines elliptischen Integrals erster Gattung ist.*

Aus Satz 4 läßt sich auf bekannte Weise verhältnismäßig leicht schließen, daß $\lambda(\tau)$ sich nach dem Spiegelungsprinzip *eindeutig* in die ganze obere τ-Halbebene fortsetzen läßt, aber nicht über die reelle Achse hinaus. *Aus Symmetriegründen folgt ferner, daß die in Fig. 4b mit I bis VI bezeichneten Bereiche einzeln auf die in Fig. 4a gleich bezeichneten Teilbereiche des Bereiches* (2.2) *abgebildet werden*[7]).

[7]) Man kann die Abbildung der *Ränder* dieser Bereiche auch leicht direkt einsehen: Falls $\left|\frac{1}{\lambda}\right| = 1$ ist, so folgt aus der Symmetrie der Verzweigungspunkte zur Geraden $\arg z = \frac{1}{2}\arg\frac{1}{\lambda} \pmod{\pi}$, daß $\left|\int_0^{1/\lambda}\right| = \left|\int_0^1\right|$ oder [vgl. Fig. 3] $|1-\tau| = 1$ ist. Ebenso folgt aus $\left|1-\frac{1}{\lambda}\right| = 1$ die Gleichung $\left|\int_1^{1/\lambda}\right| = \left|\int_0^1\right|$ oder $|\tau| = 1$ und aus $\Re\mathrm{e}\,\frac{1}{\lambda} = \frac{1}{2}$ die Gleichung $\left|\int_0^{1/\lambda}\right| = \left|\int_\lambda^{1/\lambda}\right|$, d. h. $|1-\tau| = |\tau|$ oder $\Re\mathrm{e}\,\tau = \frac{1}{2}$.

2. Perioden und Integrationswege. Die analytische Fortsetzung von $\tau(\lambda)$ läßt sich auch in der $1/\lambda$-Ebene anschaulich verfolgen. Man bewege nämlich $1/\lambda$ auf einem $[0, 1, \infty$ vermeidenden] Wege, der einem Weg von τ in der oberen τ-Halbebene entspricht und bewege die Integrationswege der K und $-iK'$ definierenden bestimmten Integrale (1.3), wenn es nötig wird, stetig mit. Aus den geradlinigen Integrationswegen können so komplizierte Wege werden. Die Integrale (1.3) ändern sich, wenn $1/\lambda$ einen geschlossenen Weg durchläuft, um Integrale längs *geschlossenen* Wegen, d.h. um *gerade* Vielfache von K und iK', was man z.B. aus Satz 3 schließen kann. Wir wollen zeigen, daß die Perioden $K(\lambda)$ und $iK'(\lambda)$ bei solchen analytischen Fortsetzungen stets *primitive* Perioden von $Z(u)$ bleiben.

Die bestimmten Integrale $K = \int_0^1$ und $iK' = \int_0^{-\infty}$ hängen nämlich analytisch von λ ab, wenn der vierte Verzweigungspunkt $1/\lambda$ innerhalb der längs der

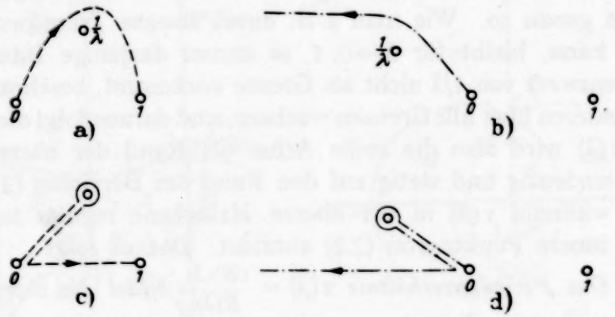


Fig. 5a—d

Halbgeraden $[-\infty, 1]$ aufgeschnittenen Ebene bewegt wird. Beim Überschreiten dieser Halbgeraden im Intervall $(0, 1)$ bzw. $(-\infty, 0)$ unter Vermeidung der Punkte $0, 1, \infty$ — etwa von oben nach unten — werden $K(\lambda)$ bzw. $iK'(\lambda)$ analytisch fortgesetzt, haben aber danach nicht mehr die Bedeutung als *geradlinig* genommene bestimmte Integrale (1.3). Die geradlinigen Integrale *nach* der Überschreitung seien mit $\hat{K}(\lambda)$ bzw. $i\hat{K}'(\lambda)$ bezeichnet. Vor der Überschreitung entsprechen ihnen Integrale längs der in Fig. 5a und b eingezeichneten Integrationswege, die man in die Wege der Fig. 5c und d deformieren kann, ohne an den Integralwerten etwas zu ändern. Diese Wege werden nach Satz 2 auf Wege in der w-Ebene abgebildet, die man in Fig. 3 bequem verfolgen kann (dort gestrichelt bzw. strichpunktiert). Daraus folgt bei Überschreitung des Intervalls $[0, 1]$

$$i\hat{K}' = iK', \qquad \hat{K} = 2iK' + K,$$

bei Überschreitung des Intervalls $(-\infty, 0)$

$$i\hat{K}' = iK' + 2K, \qquad \hat{K} = K.$$

Das sind unimodulare Substitutionen. Durchläuft also $1/\lambda$ irgendeinen geschlossenen Weg, der $0, 1, \infty$ vermeidet, so läuft die analytische Fortsetzung der Halbperioden $K(\lambda)$ und $iK'(\lambda)$ auf eine Folge unimodularer Substitutionen, d.h. auf eine unimodulare Substitution hinaus, die übrigens mod 2 der Identität kongruent ist, w. z. b. w. Bei diesen Substitutionen gehen alle Halbperioden in ihnen modulo ganzer Perioden kongruente Halbperioden über, d.h. die Pole der Funktion $Z(w)$ vertauschen sich nur unter sich, und dasselbe gilt für die 1-Stellen sowie für die $1/\lambda$-Stellen. Aus den bisher bewiesenen Tatsachen folgt unmittelbar

Satz 5. *Zu beliebig gegebenen komplexen Zahlen*[8]) ω_c, ω_n *mit* $\mathfrak{Im}\,\frac{\omega_n}{\omega_c} > 0$ *gibt es komplexe Zahlen* $\lambda \neq 0, 1$ *und* $\mu \neq 0$, *so daß die Umkehrfunktion des elliptischen Integrals 1. Gattung*

$$u = \frac{1}{\mu}\int_0^Z \frac{dZ}{2\sqrt{Z(1-Z)(1-\lambda Z)}} \tag{2.5}$$

eine doppeltperiodische meromorphe Funktion wird, die $2\omega_c$ *und* $2\omega_n$ *als primitive Perioden hat und die bei* $u \equiv \omega_n \pmod{2\omega_c, 2\omega_n}$ *Pole, bei* $u \equiv 0 \pmod{2\omega_c, 2\omega_n}$ *Nullstellen — beide von zweiter Ordnung — hat, und sonst die Werte* 0 *und* ∞ *nirgends annimmt. Dabei ist*

$$\mu\,\omega_c = \int_0^1, \qquad \mu\,\omega_n = \int_0^{-\infty}, \tag{2.6}$$

wobei die Integrationswege dadurch entstehen, daß der Verzweigungspunkt $1/\lambda$ *einen geschlossenen Weg durchläuft und die Integrationswege — ausgehend von geradlinigen Wegen — stetig mitgeführt werden.*

Der Weg des Verzweigungspunktes $1/\lambda$ bildet sich dabei in der τ-Ebene auf einen Weg ab, der den Punkt ω_n/ω_c mit einem äquivalenten [hinsichtlich der Gruppe der von einer geraden Anzahl von Spiegelungen an den Randbögen von (2.2) erzeugten linearen Transformationen] verbindet. Durch konforme Abbildung eines beliebigen solchen Weges mit Hilfe der Funktion $1/\lambda(\tau)$ [nach Satz 4 und seinen Folgerungen] erhält man stets einen geschlossenen Weg für $1/\lambda$, der nach der Vorschrift des Satzes 5 auf Integrationswege des Integrals (2.5) führt, für die die Beziehungen (2.6) gelten.

§ 3. Jacobische elliptische Funktionen

1. Grundeigenschaften. Ab jetzt werden die Liouvilleschen Sätze dauernd verwendet. Für die Umkehrfunktion $Z(u)$ des elliptischen Integrals (1.2) verwenden wir die klassische Jacobi-Gudermannsche Bezeichnung

$$Z(u) = \mathrm{sn}^2(u; k) \tag{3.1}$$

mit $k^2 = \lambda$.

Wie schon in der Einleitung angekündigt, wollen wir die Bezeichnung so abändern, daß die Perioden der Funktion $\mathrm{sn}^2 u$ beliebig vorgegeben werden

[8]) Die Bezeichnungen ω_c, ω_n werden mit Rücksicht auf das Folgende eingeführt; vgl. NEVILLE p. 51, wo statt dessen ω_1 und ω_3 geschrieben wird.

können. Das gelingt mit Hilfe der Festsetzung

$$\text{(3.2)} \qquad \mathrm{sn}^2(u;\lambda,\mu) = \frac{1}{\mu^2}\,\mathrm{sn}^2(\mu u; k).$$

Dann wird $\mathrm{sn}^2(u;\lambda,\mu)$ Umkehrfunktion des elliptischen Integrals

$$\text{(3.3)} \qquad u = \int \frac{dt}{2\sqrt{t(1-\mu^2 t)(1-\lambda\mu^2 t)}}.$$

Primitive Halbperioden werden — bei geradlinigen Integrationswegen —

$$\text{(3.4)} \qquad \omega = \int_0^{1/\mu^2}, \qquad \omega' = \int_0^{-\infty}.$$

Dabei liegt $\tau = \frac{\omega'}{\omega}$ für $\Im\mathfrak{m}\,\lambda \geq 0$ in der abgeschlossenen Hülle des Bereichs (2.2) und ist von $0, 1, \infty$ verschieden. Es ist offenbar

$$\text{(3.5)} \qquad \omega(\lambda,\mu) = \frac{K(k)}{\mu}, \qquad \omega'(\lambda,\mu) = \frac{K'(k)}{\mu}.$$

Sind ω_c, ω_n zu ω, ω' äquivalente Halbperioden, so daß $\omega_c - \omega$ und $\omega_n - \omega'$ *ganze* Perioden sind, so können wir nach Satz 5 $\mathrm{sn}^2(u;\lambda,\mu)$ auch wie folgt kennzeichnen:

Satz 6. $\mathrm{sn}^2(u;\lambda,\mu)$ *ist diejenige elliptische Funktion mit den primitiven Perioden* $2\omega_c, 2\omega_n$, *die bei* $u \equiv 0 \pmod{2\omega_c, 2\omega_n}$ *doppelte Nullstellen, bei* $u \equiv \omega_n \pmod{2\omega_c, 2\omega_n}$ *doppelte Pole hat und sonst die Werte* 0 *und* ∞ *nicht annimmt und deren Entwicklung im Nullpunkt mit dem Glied* u^2 *beginnt.*

Natürlich hätte man die Existenz dieser Funktion auch nach Weierstraßschem Vorbild mittels einer Partialbruchreihe zeigen können und dann daraus die Differentialgleichung der Funktion [vgl. (3.21)] ableiten können. Unser Ergebnis zeigt, daß wir schreiben können

$$\text{(3.6)} \qquad \begin{cases} \mathrm{sn}^2(u;\lambda,\mu) = \mathrm{sn}^2(u|\omega_c,\omega_n) = \mathrm{sn}^2(u|\omega,\omega'), \\ \mathrm{sn}^2(\varrho u|\varrho\omega_c,\varrho\omega_n) = \varrho^2\cdot\mathrm{sn}^2(u|\omega_c,\omega_n) \qquad \text{für alle } \varrho. \end{cases}$$

Man kann also $\mathrm{sn}^2 u$ ebensogut als Funktion von u, λ, μ wie als Funktion von u, ω_c, ω_n betrachten; wie bei den Weierstraßschen Funktionen. Wir werden die Parameter λ, μ bzw. ω_c, ω_n in den Funktionsbezeichnungen weglassen, wenn keine Gefahr der Verwechslung besteht. Wenn sich bei dieser Schreibweise Formeln ergeben, die von der klassischen Theorie abweichen, so tritt in ihnen μ auf, und sie gehen in die üblichen Formeln über, wenn man $\mu = 1$ setzt. Setzen wir noch

$$\text{(3.7)} \qquad \omega_d = -\omega_c - \omega_n,$$

so wird nach Satz 3 und der Definitionsgleichung (3.2)

$$\text{(3.8)} \qquad \mathrm{sn}^2\omega_c = \frac{1}{\mu^2}, \qquad \mathrm{sn}^2\omega_d = \frac{1}{\lambda\mu^2}.$$

Wie in der Weierstraßschen Theorie zeigt man leicht, daß jede gerade elliptische Funktion mit den Perioden $2\omega_c, 2\omega_n$ eine rationale Funktion von

$\mathrm{sn}^2 u$ ist und daß daher der Körper der elliptischen Funktionen mit diesen Perioden von der Funktion $\mathrm{sn}^2 u$ und ihrer Ableitung erzeugt wird.

Das von den primitiven Perioden $2\omega_c$, $2\omega_n$ der Funktion $\mathrm{sn}^2 u$ erzeugte Gitter sei ein für allemal mit $\mathfrak{W}$ bezeichnet.

Wir definieren nun

$$\mathrm{sn}^2(u|\omega_d, \omega_c) = \mathrm{sc}^2(u|\omega_c, \omega_n) = \mathrm{sc}^2 u, \tag{3.9}$$

$$\mathrm{sn}^2(u|\omega_n, \omega_d) = \mathrm{sd}^2(u|\omega_c, \omega_n) = \mathrm{sd}^2 u. \tag{3.10}$$

Die Nullstellen und Pole der Funktionen $\mathrm{sn}^2 u$, $\mathrm{sc}^2 u$, $\mathrm{sd}^2 u$ sind von zweiter Ordnung, daher sind die Quadratwurzeln $\mathrm{sn}\, u$, $\mathrm{sc}\, u$, $\mathrm{sd}\, u$ aus diesen Funktionen *eindeutige, meromorphe und ungerade Funktionen* mit den (nicht primitiven) Perioden $4\omega_c$, $4\omega_n$. Die Vorzeichen der Quadratwurzeln seien so gewählt, daß ihre Taylorschen Entwicklungen im Nullpunkt mit dem Glied u beginnen. Wir setzen weiter nach GLAISHER (vgl. NEVILLE p. 177)

$$\mathrm{cn}\, u = \frac{\mathrm{sn}\, u}{\mathrm{sc}\, u}, \qquad \mathrm{dn}\, u = \frac{\mathrm{sn}\, u}{\mathrm{sd}\, u}, \qquad \mathrm{cd}\, u = \frac{\mathrm{sd}\, u}{\mathrm{sc}\, u}. \tag{3.11}$$

Sind ferner p, q irgend zwei der vier Buchstaben s, n, c, d, so definieren wir sechs weitere Funktionen durch

$$\mathrm{qp}\, u = \frac{1}{\mathrm{pq}\, u}. \tag{3.12}$$

Schreiben wir der Einfachheit halber für 0 umständlich ω_s, so können wir die grundlegenden Eigenschaften der damit eingeführten 12 Jacobischen Funktionen wie folgt zusammenfassen:

Satz 7. *Schreibt man zwei verschiedene unter den vier Buchstaben* c, d, n, s [*etwa* p, q] *hintereinander, so erhält man ein Glaishersches Funktionssymbol für eine elliptische Funktion* pq u *mit den Perioden $4\omega_c$, $4\omega_n$.* pq u *hat Nullstellen 1. Ordnung bei ω_p (mod $\mathfrak{W}$), Pole 1. Ordnung bei ω_q (mod $\mathfrak{W}$) und in der Umgebung des Nullpunktes eine mit einem der Glieder* 1, u *oder u^{-1} beginnende Reihenentwicklung nach steigenden Potenzen von u.* pq u *ist eine ungerade oder gerade Funktion von u, je nachdem ob der Buchstabe* s *unter den Buchstaben* p, q *vorkommt oder nicht.*

Im ersten Fall können wir unsere Funktion mit sp u *oder* ps u *bezeichnen, und $2\omega_p$ ist dann unter den 3 Zahlen $2\omega_c$, $2\omega_n$, $2\omega_d$ ihre einzige Periode. Im zweiten Fall sei* r *der dritte Buchstabe aus dem Tripel* c, d, n (*außer* p *und* q). *Dann ist $2\omega_r$ Periode von* pq u, *während $2\omega_p$ und $2\omega_q$ keine Perioden sind.*

Beweis. Die Aussagen über Nullstellen und Pole folgen unmittelbar aus den Definitionen. Keine unserer 12 Funktionen kann zwei der Zahlen $2\omega_c$, $2\omega_d$, $2\omega_n$ als Perioden haben; denn dann wären diese *drei* Zahlen Perioden und unsere Funktion wäre eine elliptische Funktion 1. Ordnung, was nicht geht. Nun ist $\mathrm{pq}(u + 2\omega_i) = \pm \mathrm{pq}\, u$ $(i = c, d, n)$. In dieser Relation gilt also das $+$-Zeichen für höchstens ein i, aber wegen (3.7) auch für mindestens ein i. Ist pq u ungerade, so kann das $+$-Zeichen nur gelten, wenn pq ω_i einen der

Werte $0, \infty$ annimmt. Ist dagegen $\mathrm{pq}\, u$ gerade, so muß das $+$-Zeichen gelten, wenn $\mathrm{pq}\, \omega_i$ von 0 und ∞ verschieden ist. Daraus folgt die Behauptung.

p, q, r *seien die drei Buchstaben* c, n, d *in irgendeiner Reihenfolge. Dann gelten die Identitäten* (Beweis nach den Liouvilleschen Sätzen)

$$\mathrm{sp}(u+\omega_p) = -\,\mathrm{sp}\,\omega_q\,\mathrm{sp}\,\omega_r\,\mathrm{ps}\,u, \tag{3.13}$$

$$\mathrm{sp}(u+\omega_q) = \mathrm{sp}\,\omega_q\,\mathrm{qr}\,u, \tag{3.14}$$

$$\mathrm{pq}(u+\omega_p) = -\,\mathrm{rs}\,\omega_p\,\mathrm{sr}\,u, \tag{3.15}$$

$$\mathrm{pq}(u+\omega_q) = -\,\mathrm{sr}\,\omega_q\,\mathrm{rs}\,u, \tag{3.16}$$

$$\mathrm{pq}(u+\omega_r) = \mathrm{pq}\,\omega_r\,\mathrm{qp}\,u. \tag{3.17}$$

Die Funktionen $\mathrm{rs}^2 u - \mathrm{qs}^2 u$ *sind frei von Polen 2. Ordnung; weil sie gerade sind, aber auch frei von Polen 1. Ordnung und daher konstant.* Es folgt

$$\mathrm{rs}^2\omega_q - \mathrm{qs}^2\omega_q = \mathrm{rs}^2\omega_r - \mathrm{qs}^2\omega_r$$

oder

$$\mathrm{rs}^2\omega_q = -\,\mathrm{qs}^2\omega_r. \tag{3.18}$$

Ferner gelten die Ableitungsformeln

$$\mathrm{sp}'u = \mathrm{qp}\,u\,\mathrm{rp}\,u, \quad \mathrm{ps}'u = -\,\mathrm{qs}\,u\,\mathrm{rs}\,u, \quad \mathrm{pq}'u = \mathrm{ps}^2\omega_q\,\mathrm{sq}\,u\,\mathrm{rq}\,u \tag{3.19}$$

und die Identitäten

$$\mathrm{po}^2 u = 1 + \mathrm{ps}^2\omega_q\,\mathrm{sq}^2 u, \quad \mathrm{qs}^2 u = \mathrm{qs}^2\omega_p + \mathrm{ps}^2 u. \tag{3.20}$$

Beweis. Beide Seiten der Gln. (3.20) haben dieselben Pole mit denselben Hauptteilen und stimmen für $u=0$ bzw. $u=\omega_p$ überein. Die Ableitungen der 12 Jacobischen Funktionen $f(u)$ verschwinden an den Stellen ω_i $(i = s, n, c, d)$, wo die Funktionen selbst $\neq \infty, 0$ sind. Denn dann ist $f'(u+2\omega_i) = \pm f'(u)$, je nachdem ob $f(u)$ gerade oder ungerade, $f'(u)$ also ungerade oder gerade ist. Setzt man $u = -\omega_i$, so ergibt sich die Behauptung. Beide Seiten der Gln. (3.19) stimmen also in ihren Nullstellen und Polen überein. Die Konstantenbestimmung in der dritten Gleichung erfolgt am einfachsten, indem man $\mathrm{pq}\,u = \mathrm{ps}\,u\,\mathrm{sq}\,u$ differenziert und (3.20) verwendet. Es folgt

$$\begin{cases} \mathrm{sp}'^2 u = (1 + \mathrm{qs}^2\omega_p\,\mathrm{sp}^2 u)\,(1 + \mathrm{rs}^2\omega_p\,\mathrm{sp}^2 u), \\ \mathrm{ps}'^2 u = (\mathrm{qs}^2\omega_p + \mathrm{ps}^2 u)\,(\mathrm{rs}^2\omega_p + \mathrm{ps}^2 u), \end{cases} \tag{3.21}$$

$$\mathrm{qp}^2\omega_r + \mathrm{qr}^2\omega_p = \frac{\mathrm{sp}^2\omega_r}{\mathrm{sq}^2\omega_r} + \frac{\mathrm{sr}^2\omega_p}{\mathrm{sq}^2\omega_p} = \mathrm{sp}^2\omega_r(\mathrm{qs}^2\omega_r - \mathrm{qs}^2\omega_p) = 1 \tag{3.22}$$

[nach (3.18) und (3.20)].

Das Produkt $\mathrm{ps}\,u\,\mathrm{ps}(u+\omega_p)$ ist nach (3.13) konstant. Entwicklung um $u=0$ ergibt als Wert dieser Konstanten $\mathrm{ps}'\omega_p$. Nach (3.13) und (3.19) ist daher

$$\mathrm{ps}\,\omega_q\,\mathrm{ps}\,\omega_r = \mathrm{qs}\,\omega_p\,\mathrm{rs}\,\omega_p$$

oder [wegen (3.18)]

$$\frac{\mathrm{sp}\,\omega_q}{\mathrm{sq}\,\omega_p} = \frac{\mathrm{sq}\,\omega_r}{\mathrm{sr}\,\omega_q} = \frac{\mathrm{sr}\,\omega_p}{\mathrm{sp}\,\omega_r} = \pm i. \tag{3.23}$$

Das Vorzeichen rechts hängt von der Wahl der primitiven Halbperioden [der Funktion $\mathrm{sn}^2 u$] $\omega_p, \omega_q, \omega_r$ ab. Aus Stetigkeitsgründen muß es fest bleiben, solange sich diese Halbperioden stetig ändern und ihre Verhältnisse nicht reell werden. Das ist der Fall, solange das in § 2 behandelte Periodenverhältnis τ in der oberen Halbebene bleibt. Für $\frac{1}{\lambda} > 1$ wird

$$\mathrm{sn}(K|K, iK') = \mathrm{sn}(K; k) = +1,$$
$$\mathrm{sc}(iK'|K, iK') = \mathrm{sn}(iK'|iK', -K) = i\,\mathrm{sn}(K'|K', iK).$$

In dem Sonderfall $K = K'$, $K = \omega_c$, $iK' = \omega_n$ ist also $\frac{\mathrm{sc}\,\omega_n}{\mathrm{sn}\,\omega_c} = +i$, und daraus folgt [NEVILLE p. 59]:

In (3.23) *gilt das* $+$*-Zeichen, wenn und nur wenn* $\Im\mathrm{m}\frac{\omega_q}{\omega_p} > 0$ *ist.*

Wir können jetzt die Vorzeichen der Quadratwurzeln $\sqrt{\mu^2} = \mu$, $\sqrt{\lambda} = k$ und $\sqrt{1-\lambda} = k'$ festlegen durch [vgl. (3.8)]

$$(3.24) \qquad \mu = \mathrm{ns}\,\omega_c, \qquad k = -\frac{\mathrm{sn}\,\omega_c}{\mathrm{sn}\,\omega_d} = \frac{i\,\mathrm{sc}\,\omega_n}{i\,\mathrm{sd}\,\omega_n} = \mathrm{dc}\,\omega_n, \qquad k' = \mathrm{dn}\,\omega_c.$$

2. Lineare Transformationen. Nach (3.23) lassen sich alle in den Gln. (3.13) bis (3.22) vorkommenden Konstanten auf bequeme Art durch die drei Größen μ, k und k' rational ausdrücken. Damit kann man die Transformationsformeln für diese drei Größen und damit für die klassischen Jacobischen Funktionen [als Funktionen von u und λ] bei *Gittertransformationen* leicht angeben, z. B.

$$k(-\omega_n, \omega_c) = \mathrm{dc}(\omega_c|-\omega_n, \omega_c) = \mathrm{dn}(\omega_c|\omega_c, \omega_n) = k',$$
$$\mu(-\omega_n, \omega_c) = \mathrm{ns}(-\omega_n|-\omega_n, \omega_c) = -\mathrm{cs}\,\omega_n = i\,\mu.$$

Daraus folgt nach (3.6) und (3.2) die Jacobische Relation

$$i\,\mathrm{sn}(u; k') = \mathrm{sn}(iu; k).$$

Das Verhalten bei Gittertransformationen zeigt besonders eindringlich, daß es zweckmäßig ist, die Jacobischen Funktionen wie die Weierstraßschen als Funktionen der Perioden zu betrachten. Die zwölf Jacobischen Funktionen vertauschen sich bei Gittertransformationen wie folgt: Bei der Substitution

$$(3.25) \qquad \begin{cases} \dot\omega_n = \alpha\omega_n + \beta\omega_c, \\ \dot\omega_c = \gamma\omega_n + \delta\omega_c, \end{cases} \qquad \alpha, \beta, \gamma, \delta \text{ ganzzahlig}, \ \alpha\delta - \beta\gamma = 1,$$

gibt es eine eindeutig bestimmte Permutation p, q, r der Buchstaben c, n, d, so daß

$$\left.\begin{aligned} \dot\omega_c &\equiv \omega_p, \\ \dot\omega_n &\equiv \omega_q, \\ \dot\omega_d &\equiv \omega_r, \end{aligned}\right\} \quad (\mathrm{mod}\,\mathfrak{W})$$

wird. Dann ist einfach

$$(3.26) \qquad \begin{cases} \dot{\mathrm{sn}}\,u = \mathrm{sn}(u|\dot\omega_c, \dot\omega_n) = \mathrm{sq}\,u, \\ \dot{\mathrm{sc}}\,u = \mathrm{sc}(u|\dot\omega_c, \dot\omega_n) = \mathrm{sp}\,u, \\ \dot{\mathrm{sd}}\,u = \mathrm{sd}(u|\dot\omega_c, \dot\omega_n) = \mathrm{sr}\,u \end{cases} \qquad \text{usw.}$$

3. Additionstheoreme. Sind p, q zwei verschiedene der vier Buchstaben s, c, n, d und ist $2\omega_r$ Periode von $\mathrm{pq}\,u$ [Satz 7], so ist bei festgehaltenem $v \not\equiv \omega_i \pmod{\mathfrak{P}}$, $i = s, c, n, d$,

$$f(u) = \frac{\mathrm{pq}\,u\,\mathrm{pq}'v + \mathrm{pq}'u\,\mathrm{pq}\,v}{\mathrm{pq}^2 u - \mathrm{pq}^2 v}$$

eine elliptische Funktion mit den Perioden $4\omega_c$, $4\omega_n$, $2\omega_r$ und der Eigenschaft

$$f(u + 2\omega_i) = \pm f(u), \qquad i = c, n, d.$$

Sie ist bei $u \equiv -v$ und $u \equiv \omega_q \pmod{\mathfrak{P}}$ regulär und hat daher nur bei $u \equiv v$ Pole. Bei $u \equiv v + \omega_r$ hat sie nach (3.17) bzw. (3.13) Nullstellen. Daher ist

$$f(u) = a\,\mathrm{rs}(u - v)$$

mit von u — und aus Symmetriegründen auch von v — unabhängigem a. Das gilt auch noch für $v = 0$. So ergibt sich im Fall $\mathrm{p}, \mathrm{q} \neq \mathrm{s}$

$$a\,\mathrm{rs}\,u = \frac{\mathrm{pq}'u}{\mathrm{pq}^2 u - 1} = \mathrm{rs}\,u \quad \text{[nach (3.19) und (3.20)], also } a = 1.$$

Im Fall $\mathrm{p} = \mathrm{s}$, $\mathrm{q} = \mathrm{r}$ wird $a\,\mathrm{rs}\,u = \mathrm{qp}\,u$, also $a = 1$.
Im Fall $\mathrm{q} = \mathrm{s}$, $\mathrm{p} = \mathrm{r}$ wird $a\,\mathrm{rs}\,u = \mathrm{pq}\,u$, also $a = 1$.
Wir haben damit [vgl. NEVILLE S. 203]

Satz 8. *Sind* p, q, r *irgendwelche der vier Buchstaben* s, n, c, d, *so daß* $\mathrm{rs}\,u$ *und* $\mathrm{pq}\,u$ *erklärt sind und dieselben Perioden haben, so gilt das Additionstheorem*

$$\mathrm{rs}(u - v) = \frac{\mathrm{pq}\,u\,\mathrm{pq}'v + \mathrm{pq}'u\,\mathrm{pq}\,v}{\mathrm{pq}^2 u - \mathrm{pq}^2 v}. \tag{3.27}$$

Weitere Additionstheoreme können leicht nach den üblichen Methoden gefunden [NEVILLE p. 74—85 und p. 200—207] oder aus (3.27) nach (3.13) bis (3.17) abgeleitet werden.

4. Transformation zweiter Ordnung. *Es gelten die Formeln* (p, q, r = c, n, d)

$$\mathrm{ps}\,u + \mathrm{qs}\,u = 2\,\mathrm{cs}(u\,|\,2\omega_p, \omega_r), \tag{3.28}$$

$$\mathrm{ps}\,u - \mathrm{qs}\,u = 0{,}5\,\mathrm{ps}^2\omega_q\,\mathrm{sc}(u\,|\,2\omega_p, \omega_r), \tag{3.29}$$

$$\mathrm{qr}\,u + \mathrm{qp}\omega_r\,\mathrm{pr}\,u = (1 + \mathrm{qp}\omega_r)\,\mathrm{dn}(u\,|\,2\omega_p, \omega_r), \tag{3.30}$$

$$\mathrm{qr}\,u - \mathrm{qp}\omega_r\,\mathrm{pr}\,u = (1 - \mathrm{qp}\omega_r)\,\mathrm{nd}(u\,|\,2\omega_p, \omega_r). \tag{3.31}$$

Beweis. In jeder der vier Formeln haben die Funktionen auf beiden Seiten dieselben primitiven Perioden, nämlich $4\omega_c$ und $4\omega_n$, dieselben Pole und dieselben Nullstellen. Die vier Funktionen auf der linken Seite bleiben nämlich endlich für $u = 2\omega_p$ bzw. $u = 0$ bzw. $u = \omega_r + 2\omega_p$ bzw. $u = \omega_r$. Weil sie ungerade bzw. gerade sind, müssen sie dort sogar Nullstellen haben. Also sind die linken Seiten den rechten proportional. Die Proportionalitätsfaktoren ergeben sich für (3.28), (3.30) und (3.31) durch Einsetzen von $u = 0$, für (3.29) aus (3.28) und (3.20).

In dem Sonderfall $p=c$, $r=n$ erhält man daraus mit Hilfe von (3.22) und (3.23) leicht die Identitäten

$$(3.32)\qquad \mu(2\omega_c,\omega_n)=\frac{1+k}{2}\mu,\qquad k'(2\omega_c,\omega_n)=\frac{1-k}{1+k},$$

aus denen sich der GAUSSsche Satz vom arithmetisch-geometrischen Mittel ableiten läßt.

Eine besonders einfache Transformationsformel ist

$$(3.33)\qquad \mathrm{sq}^2\omega_r\frac{\mathrm{qr}'u}{\mathrm{qr}\,u}=\mathrm{sr}\,u\,\mathrm{pq}\,u=\mathrm{sn}\left(u\,\middle|\,\frac{\omega_p}{2},\omega_r\right).$$

Sie folgt aus der Gleichheit der Perioden, Nullstellen und Pole auf beiden Seiten. Ebenso ist

$$(3.34)\qquad \frac{\mathrm{sp}'u}{\mathrm{sp}\,u}=\mathrm{qs}\,u\,\mathrm{rp}\,u=\mathrm{ns}\left(u\,\middle|\,\frac{\omega_p}{2},\omega_r\right).$$

Erwähnt seien noch die damit verwandten und für die unbestimmte Integration der Jacobischen elliptischen Funktionen brauchbaren Formeln [NEVILLE S. 234]

$$(3.35)\qquad \mathrm{rs}\,u=\frac{\mathrm{ps}'u-\mathrm{qs}'u}{\mathrm{ps}\,u-\mathrm{qs}\,u}=-\frac{\mathrm{ps}'u+\mathrm{qs}'u}{\mathrm{ps}\,u+\mathrm{qs}\,u},$$

$$(3.36)\qquad \mathrm{rs}'\omega_r\,\mathrm{sr}\,u=\frac{\mathrm{qr}'u-\mathrm{qp}\,\omega_r\,\mathrm{pr}'u}{\mathrm{qr}\,u-\mathrm{qp}\,\omega_r\,\mathrm{pr}\,u}=-\frac{\mathrm{qr}'u+\mathrm{qp}\,\omega_r\,\mathrm{pr}'u}{\mathrm{qr}\,u+\mathrm{qp}\,\omega_r\,\mathrm{pr}\,u},$$

$$(3.37)\qquad \mathrm{rs}\,\omega_p\,\mathrm{qp}\,u=\frac{\mathrm{sp}'u+\mathrm{sr}\,\omega_p\,\mathrm{rp}'u}{\mathrm{sp}\,u+\mathrm{sr}\,\omega_p\,\mathrm{rp}\,u}=-\frac{\mathrm{sp}'u-\mathrm{sr}\,\omega_p\,\mathrm{rp}'u}{\mathrm{sp}\,u-\mathrm{sr}\,\omega_p\,\mathrm{sp}\,u}.$$

Man zeigt nämlich leicht, daß rechts und links die Perioden, Nullstellen und Pole übereinstimmen. Man kann auch die zweite und dritte Gleichung mittels (3.13) und (3.14) auf die erste zurückführen.

5. Weierstraßsche Funktionen. Gegen Gittertransformationen invariante Funktionen des Körpers $C(\omega_c,\omega_n)$ erhält man durch folgende *Definitionen:*

$$(3.38)\qquad \begin{cases} \wp u=\frac{1}{3}\left(\mathrm{cs}^2u+\mathrm{ns}^2u+\mathrm{ds}^2u\right),\\ \wp' u=-2\,\mathrm{cs}\,u\,\mathrm{ns}\,u\,\mathrm{ds}\,u=\frac{d}{du}\,\mathrm{ps}^2u\qquad (\mathrm{p}=\mathrm{n},\mathrm{c},\mathrm{d}).\end{cases}$$

Daß $\wp' u$ die Ableitung von $\wp u$ ist, folgt aus (3.19). $\wp u$ hat einen Pol vom Hauptteil u^{-2} an der Stelle $u=0$ und sonst (modulo Perioden) keine Pole. Nach (3.18) ist $\wp\omega_c+\wp\omega_n+\wp\omega_d=0$. Daher ist $\wp u$ mit der Weierstraßschen $\wp$-Funktion identisch.

§ 4. Jacobische Zeta- und Thetafunktionen

1. Definitionen. Das elliptische Integral 2. Gattung

$$(4.1)\qquad E(u)=\int_0^u \mathrm{dn}^2u\,du=\int_0^u\left[\left(\frac{1}{\mathrm{nd}'\omega_n(u-\omega_n)}\right)^2+\cdots\right]du=\frac{1}{\mu^2(u-\omega_n)}+\cdots$$

ist eine eindeutige, ungerade Funktion. Wir setzen

$$(4.2)\qquad \mathrm{zn}\,u=\mathrm{zn}\,(u|\omega_c,\omega_n)=\mathrm{zn}\,(u;\lambda,\mu)=\mu^2E(u)+a\cdot u$$

mit konstantem a. Die Differenzen $\mathrm{zn}(u+2\omega_c)-\mathrm{zn}\,u$ und $\mathrm{zn}(u+2\omega_n)-\mathrm{zn}\,u$ sind Konstante. Wir legen a so fest, daß die erste verschwindet. Für die zweite ergibt sich dann durch Bildung des Umlaufintegrals $\int \mathrm{zn}\,u\,du$ um ein Periodenparallelogramm, weil alle Residuen von $\mathrm{zn}\,u$ den Wert 1 haben, der Wert $\frac{-\pi i}{\omega_c}$. Es ist also

$$\text{(4.3)} \qquad \mathrm{zn}(u+2\omega_c)-\mathrm{zn}\,u=0, \qquad \mathrm{zn}(u+2\omega_n)-\mathrm{zn}\,u=\frac{-\pi i}{\omega_c}.$$

Diese Jacobische Zetafunktion leistet dasselbe wie die Weierstraßsche, nämlich (mit ihren Ableitungen) die *Partialbruchdarstellung beliebiger elliptischer Funktionen*. Die hier eingeführte Funktion $\mathrm{zn}\,u$ hängt mit der klassischen Jacobischen Zetafunktion $\mathrm{zn}(u;k)$ zusammen durch die Gleichung

$$\text{(4.4)} \qquad \mathrm{zn}(u;\lambda,\mu)=\mu\,\mathrm{zn}(\mu u;k).$$

Wie üblich setzen wir zur Abkürzung

$$\text{(4.5)} \qquad v=\frac{u}{2\omega_c}, \quad z=e^{\pi i v}, \quad \tau=\omega_n:\omega_c, \quad q=e^{\pi i \tau}, \quad (|q|<1),$$

ferner

$$\text{(4.6)} \qquad \mathrm{zn}\,u=\frac{\vartheta_n'(u)}{\vartheta_n(u)}.$$

2. Reihen- und Produktentwicklungen. Die Funktionen

$$\mu^2\,\mathrm{dn}^2 u \quad \text{und} \quad -\left(\frac{\pi}{2\omega_c}\right)^2\sum_{m=-\infty}^{\infty}\sin^{-2}\frac{\pi}{2\omega_c}\,[u-(2m-1)\,\omega_n]$$

haben dieselben Pole mit denselben Hauptteilen und unterscheiden sich daher nur um eine additive Konstante C. Die Funktion $C+\mu^2\,\mathrm{dn}^2 u$ kann nur auf eine Art so gewählt werden, daß ihr Integral $\mathrm{zn}\,u$ die Periode $2\omega_c$ hat. Weil $\mathrm{zn}\,u$ ungerade ist, ist dadurch auch die Integrationskonstante eindeutig bestimmt und es wird

$$\text{(4.7)} \qquad \mathrm{zn}\,u=\frac{\pi}{2\omega_c}\sum_{m=1}^{\infty}\left\{\mathrm{ctg}\,\frac{\pi}{2\omega_c}\,[u-(2m-1)\,\omega_n]+\mathrm{ctg}\,\frac{\pi}{2\omega_c}\,[u+(2m-1)\,\omega_n]\right\}$$

Wir setzen zur Abkürzung

$$\text{(4.8)} \qquad \begin{cases} Q_1=\prod_{m=1}^{\infty}(1-q^{2m}), & Q_2=\prod_{m=1}^{\infty}(1+q^{2m}), \\ Q_3=\prod_{m=1}^{\infty}(1+q^{2m-1}), & Q_4=\prod_{m=1}^{\infty}(1-q^{2m-1}). \end{cases}$$

Dann wird $Q_1\,Q_2\,Q_3\,Q_4=Q_1$ oder

$$\text{(4.9)} \qquad Q_2\,Q_3\,Q_4=1.$$

Aus (4.7) folgt durch Integration nach kurzer Rechnung

$$\text{(4.10)} \qquad \vartheta_n(u)=Q_4^{-2}\prod_{m=1}^{\infty}(1-z^2q^{2m-1})\,(1-z^{-2}q^{2m-1})=Q_4^{-2}F(z).$$

Dabei ist die Integrationskonstante so festgelegt, daß

$$\vartheta_n(0) = 1$$

wird. Aus der Funktionalgleichung

$$F(qz) = -q^{-1}z^{-2}F(z) \tag{4.11}$$

erhält man die Laurentsche Entwicklung

$$F(z) = A(q)\sum_{m=-\infty}^{\infty} q^{m^2}(iz)^{2m} \tag{4.12}$$

mit noch zu bestimmendem $A(q)$. Setzt man $z=1$ und $z=\exp\left(-\frac{\pi i}{4}\right)$ ein, so erhält man durch Vergleich die Identität[9])

$$A(q) : A(q^4) = \prod_{m=1}^{\infty}(1-q^{8m}) : \prod_{m=1}^{\infty}(1-q^{2m}).$$

Wegen $\lim\limits_{q\to 0} A(q) = 1$ folgt daraus $A(q) = Q_1^{-1}$. Aus (4.10), (4.11) und (4.12) folgt

$$\begin{cases} \vartheta_n(u) = \dfrac{1 - 2q\cos 2\pi v + 2q^4\cos 4\pi v - 2q^9\cos 6\pi v + \cdots}{1 - 2q + 2q^4 - 2q^9 + \cdots} \\ \qquad = Q_4^{-2}\prod\limits_{m=1}^{\infty}(1 - z^2q^{2m-1})(1 - z^{-2}q^{2m-1}). \end{cases} \tag{4.13}$$

Aus (4.13) und (4.11) folgen unmittelbar die Funktionalgleichungen

$$\vartheta_n(u + 2\omega_c) = \vartheta_n(u), \quad \vartheta_n(u + 2\omega_n) = -z^{-2}q^{-1}\vartheta_n(u). \tag{4.14}$$

Setzen wir noch [vgl. NEVILLE p. 274]

$$\vartheta_s(u) = \frac{z\vartheta_n(u+\omega_n)}{\vartheta_n'(\omega_n)} = \frac{2\omega_c}{\pi}\,\frac{\sin\pi v - q^{1\cdot 2}\sin 3\pi v + q^{2\cdot 3}\sin 5\pi v + \cdots}{1 - 3q^2 + 5q^6 - 7q^{12} + \cdots}, \tag{4.15}$$

$$\vartheta_c(u) = \frac{\vartheta_s(u+\omega_c)}{\vartheta_s(\omega_c)} = \frac{\cos\pi v + q^2\cos 3\pi v + q^6\cos 5\pi v + \cdots}{1 + q^2 + q^6 + q^{12} + q^{20} + \cdots}, \tag{4.16}$$

$$\vartheta_d(u) = \frac{\vartheta_n(u+\omega_c)}{\vartheta_n(\omega_c)} = \frac{1 + 2q\cos 2\pi v + 2q^4\cos 4\pi v + \cdots}{1 + 2q + 2q^4 + 2q^9 + \cdots}, \tag{4.17}$$

so erhalten wir für $p = s, n, c, d$

$$\vartheta_p(u + 2\omega_c) = \pm\vartheta_p(u), \tag{4.18}$$

$$\vartheta_p(u + 2\omega_n) = \pm q^{-1}z^{-2}\vartheta_p(u). \tag{4.19}$$

Das +-Zeichen gilt in (4.18) für $p = n, d$, in (4.19) für $p = c, d$. Die Funktion $\vartheta_p(u)$ ist ungerade für $p = s$ und sonst gerade.

3. Jacobische elliptische Funktionen als Thetaquotienten. Es gilt die zusammenfassende Formel [für p, q = s, c, n, d]

$$\frac{\vartheta_p(u)}{\vartheta_q(u)} = \mathrm{pq}(u) \quad [\text{NEVILLE p. 275}]. \tag{4.20}$$

[9]) Vgl. H. LENZ, Jahresber. DMV 58 (1956) S. 57.

Die in (4.15) bis (4.17) vorkommenden Konstanten $\vartheta_p(\omega_q)$ lassen sich leicht bestimmen. Es wird, wenn wir abkürzend $q_0 = q^{-\frac{1}{4}}$ setzen,

$$(4.21)\quad \begin{cases} \vartheta_n'(\omega_n) = i\mu\sqrt{k}\,q_0, & \vartheta_s(\omega_n) = \dfrac{i\,q_0}{\mu\sqrt{k}}, \\ \vartheta_c(\omega_n) = \dfrac{q_0}{\sqrt{k}}, & \vartheta_d(\omega_n) = q_0\sqrt{k}\,; \end{cases}$$

$$(4.22)\quad \begin{cases} \vartheta_c'(\omega_c) = -\mu\sqrt{k'}, & \vartheta_n(\omega_c) = \dfrac{1}{\sqrt{k'}}, \\ \vartheta_s(\omega_c) = \dfrac{1}{\mu\sqrt{k'}}, & \vartheta_d(\omega_c) = \sqrt{k'}\,; \end{cases}$$

$$(4.23)\quad \begin{cases} \vartheta_d'(\omega_d) = -i\mu q_0\sqrt{k\,k'}, & \vartheta_c(\omega_d) = -i\,q_0\sqrt{\dfrac{k'}{k}}, \\ \vartheta_n(\omega_d) = q_0\sqrt{\dfrac{k}{k'}}, & \vartheta_s(\omega_d) = \dfrac{-q_0}{\mu\sqrt{k\,k'}}. \end{cases}$$

Man erhält nämlich aus (4.15) für $u = \omega_n$

$$\vartheta_s(\omega_n)\,\vartheta_n'(\omega_n) = -q^{-\frac{1}{2}},$$

andererseits aus (4.20), (3.19), (3.23) und (3.24) nach kurzer Rechnung

$$\frac{\vartheta_n'(\omega_n)}{\vartheta_s(\omega_n)} = \mu^2 k.$$

Daraus und aus (4.20) folgt (4.21), wobei das noch nicht festgelegte Vorzeichen von $\sqrt{k}$ so zu bestimmen ist, daß es für $\Re\mathrm{e}\,\tau = 0$ positiv wird (dann ist $0 < k < 1$). Entsprechend leitet man (4.22) ab. (4.23) folgt aus (4.21) und (4.22), wenn man (4.15) bis (4.17) verwendet. Aus (4.22) folgt die bekannte Formel

$$(4.24)\qquad \sqrt{k'} = \frac{1 - 2q + 2q^4 - 2q^9 + \cdots}{1 + 2q + 2q^4 + 2q^9 + \cdots},$$

aus der man die für die numerische Rechnung wichtige Entwicklung

$$q = \gamma + 2\gamma^5 + 15\gamma^9 + 150\gamma^{13} + \cdots, \qquad 2\gamma = \frac{1 - \sqrt{k'}}{1 + \sqrt{k'}}$$

ableitet[10]).

Aus (4.13) ergibt sich

$$\vartheta_n(\omega_c) = Q_4^{-2}\,Q_3^2, \qquad \vartheta_n(\omega_d) = 2\,Q_4^{-2}\,Q_2^2, \qquad \vartheta_n'(\omega_n) = \frac{\pi i}{\omega_c}\,Q_4^{-2}\,Q_1^2.$$

Daher wird, wenn man (4.9) berücksichtigt,

$$i\mu = \frac{\vartheta_n'(\omega_n)\,\vartheta_n(\omega_c)}{\vartheta_n(\omega_d)} = \frac{\pi i}{2\omega_c}\,Q_1^2\,Q_3^4.$$

---

[10]) TRICOMI, F.: Elliptische Funktionen, S. 176. Leipzig 1948.

Setzt man in (4.12) $z=i$, so wird wegen $A(q)=Q_1^{-1}$

$$Q_1 Q_3^2 = 1 + 2q + 2q^4 + 2q^9 + \cdots \tag{4.25}$$

und daher [TRICOMI S. 175]

$$\sqrt{\frac{2\mu\omega_c}{\pi}} = 1 + 2q + 2q^4 + \cdots,$$

nach (4.24) also

$$\sqrt{\frac{\mu\omega_c}{2\pi}}\left(1+\sqrt{k'}\right) = 1 + 2q^4 + 2q^{16} + 2q^{36} + \cdots. \tag{4.26}$$

4. Lineare Transformation der Thetafunktionen. Wir wollen den Einfluß der Periodensubstitutionen

$$\begin{cases} \dot\omega_n = \alpha\omega_n + \beta\omega_c, \\ \dot\omega_c = \gamma\omega_n + \delta\omega_c \end{cases} \quad \text{mit } \alpha\delta - \beta\gamma = 1 \tag{3.25}$$

auf die NEVILLEschen Thetafunktionen $\vartheta_p(u)$ untersuchen. Es sei für $m, p = n, c, d, s$

$$\dot\omega_m \equiv \omega_p \pmod{\mathfrak{W}}. \tag{4.27}$$

Wir setzen

$$\dot\vartheta_m(u) = \vartheta_m(u\,|\dot\omega_c, \dot\omega_n). \tag{4.28}$$

Dann ist

$$\frac{\dot\vartheta_m(u)}{\vartheta_p(u)} = e^{G(u)}$$

eine ganze, gerade Funktion ohne Nullstellen. Also ist auch $G(u)$ eine ganze, gerade Funktion. Setzen wir für p = s, c, n, d [vgl. (4.6)]

$$\mathrm{zp}\,u = \frac{\vartheta_p'(u)}{\vartheta_p(u)}, \qquad \dot{\mathrm{z}}\mathrm{m}\,u = \mathrm{zm}\,(u\,|\dot\omega_c, \dot\omega_n), \tag{4.29}$$

so wird nach (4.18) und (4.19) [vgl. (4.3)]

$$\mathrm{zp}\,(u+2\omega_c) - \mathrm{zp}\,u = 0, \qquad \mathrm{zp}\,(u+2\omega_n) - \mathrm{zp}\,u = \frac{-\pi i}{\omega_c}. \tag{4.30}$$

Wegen $G'(u) = \dot{\mathrm{z}}\mathrm{m}\,u - \mathrm{zp}\,u$ ist $G''(u)$ eine ganze elliptische Funktion, also eine Konstante A. Es ergibt sich

$$2\dot\omega_c A = G'(u+2\dot\omega_c) - G'(u) = \frac{\gamma\pi i}{\omega_c}.$$

Weil $G'(u)$ ungerade ist und wegen $\frac{\dot\vartheta_m(0)}{\vartheta_p(0)} = 1$ ergibt sich

$$\dot\vartheta_m(u) = \vartheta_m(u\,|\dot\omega_c, \dot\omega_n) = \exp\left(\frac{\gamma\pi i u^2}{4\omega_c\dot\omega_c}\right)\vartheta_p(u). \tag{4.31}$$

Es treten also keine achten Einheitswurzeln auf wie bei den üblichen Thetafunktionen. *Aus diesem Grunde sind die* NEVILLE*schen Thetafunktionen den*

klassischen vorzuziehen, z.B. für die Produktdarstellung einer elliptischen Funktion $f(u)$ mit den Nullstellen a_i und den Polen b_i in der Form

$$f(u) = \text{const} \cdot \prod_i \frac{\vartheta_s(u-a_i)}{\vartheta_s(u-b_i)}, \qquad \sum a_i = \sum b_i. \tag{4.32}$$

Die WEIERSTRASSsche σ-Funktion ist für die meisten Zwecke entbehrlich. Andererseits kann sie leicht durch folgenden Ansatz gewonnen werden: Man setze $\zeta(u)$ gleich der ungeraden Funktion, deren negative Ableitung die durch (3.38) erklärte Funktion $\wp(u)$ ist, ferner

$$\zeta(u) = \frac{\sigma'(u)}{\sigma(u)}, \qquad \sigma'(0) = 1, \qquad \zeta(u+2\omega_c) - \zeta(u) = 2\eta_c.$$

Dann wird

$$\sigma(u) = \exp\left(\frac{\eta_c u^2}{2\omega_c}\right)\vartheta_s(u). \tag{4.33}$$

5. Zetafunktionen. Die Funktion $\mathrm{zs}\,u$ kann genau so wie die WEIERSTRASSsche Zetafunktion zur Darstellung elliptischer Funktionen mit gegebenen Polen von gegebenen Hauptteilen dienen. Aus (4.15) bis (4.17) folgt

$$\begin{cases} \mathrm{zs}\,u = \dfrac{\pi i}{2\omega_c} + \mathrm{zn}\,(u+\omega_n), \\ \mathrm{zc}\,u = \mathrm{zs}\,(u+\omega_c), \\ \mathrm{zd}\,u = \mathrm{zn}\,(u+\omega_c). \end{cases} \tag{4.34}$$

Die vier Zetafunktionen $\mathrm{zp}\,u$ sind ungerade Funktionen, daher ist

$$\mathrm{zp}\,\omega_c = 0 \qquad \text{für } \mathrm{p} = \mathrm{s, n, d, und ebenso} \tag{4.35}$$

$$\mathrm{zp}\,\omega_n = \frac{1}{2}\left[\mathrm{zp}\,\omega_n - \mathrm{zp}(-\omega_n)\right] = \frac{-\pi i}{2\omega_c} \qquad \text{für } \mathrm{p} = \mathrm{s, c, d}; \tag{4.36}$$

$$\mathrm{zp}\,\omega_d = \frac{\pi i}{2\omega_c} \qquad \text{für } \mathrm{p} = \mathrm{s, n, c}. \tag{4.37}$$

Damit sind die Werte der ersten Ableitungen der NEVILLEschen Thetafunktionen für Halbperiodenwerte ω_p von u bekannt. Die Werte der höheren Ableitungen an diesen Stellen erhält man mit Hilfe der Formel

$$\mathrm{zn}'\,u = \mathrm{zc}'(u \pm \omega_d) = \mathrm{zs}'(u \pm \omega_n) = \mathrm{zd}'(u \pm \omega_c) = \mu^2\left[\mathrm{dn}^2 u - \frac{E(\omega_c)}{\omega_c}\right], \tag{4.38}$$

die sich aus (4.2) ergibt, wenn man die Konstante a durch $\mathrm{zn}\,\omega_c = 0$ bestimmt und differenziert.

§ 5. Eine Verallgemeinerung der Jacobischen elliptischen Funktionen

1. Funktionen mit doppeltperiodischem absolutem Betrag. Die absoluten Beträge aller 12 Jacobischen Funktionen $\mathrm{pq}\,u$ haben $\mathfrak{W}$ als Periodengitter. Es liegt daher nahe, sie als Sonderfälle der meromorphen Funktionen mit doppeltperiodischem absolutem Betrag zu betrachten. Für diese gilt der einfache

Satz 9. *Eine meromorphe Funktion $f(u) \not\equiv 0$ mit doppeltperiodischem absolutem Betrag hat ebenso viele Nullstellen wie Pole. Zu m gegebenen Nullstellen und m gegebenen Polstellen gibt es bis auf einen konstanten Faktor genau eine derartige Funktion.*

Beweis. I. Sind $f(u)$ und $g(u)$ zwei solche Funktionen und ist jede Nullstelle von $g(u)$ auch Nullstelle von $f(u)$, jeder Pol von $f(u)$ auch Pol von $g(u)$, so ist ihr Quotient $f(u):g(u)$ von beschränktem absolutem Betrag, also konstant.

II. Daß die Zahl der Nullstellen gleich der der Pole ist, folgt wie bei den elliptischen Funktionen durch Bildung des Umlaufintegrals $\int \frac{f'(u)}{f(u)}\,du$.

III. Weil die von $f(u) \equiv 0$ verschiedenen Funktionen mit doppeltperiodischem absolutem Betrag zum Periodengitter $\mathfrak{W}$ eine Gruppe hinsichtlich der Multiplikation bilden, genügt es, die Existenzaussage des Satzes 11 für $m = 1$ zu beweisen.

Es seien ω_e und $\omega_f \not\equiv \omega_e \pmod{\mathfrak{W}}$ beliebige komplexe Zahlen (also im allgemeinen *keine* Halbperioden). *Wir erklären eine Funktion* ef u *durch*

$$\text{ef}\,u = A \cdot \frac{\vartheta_s(u-\omega_e)}{\vartheta_s(u-\omega_f)}\, e^{Bu} \tag{5.1}$$

mit

$$B = \frac{\pi i}{2\omega_c} \cdot \frac{(\omega_f-\omega_e)\,\overline{\omega}_c - (\overline{\omega}_f-\overline{\omega}_e)\,\omega_c}{\omega_n\,\overline{\omega}_c - \overline{\omega}_n\,\omega_c}. \tag{5.2}$$

Die Konstante A sei so bestimmt, daß entweder ef 0 *oder* ef'0 *oder* fe'0 $\left[\text{mit } \text{fe}\,u = \frac{1}{\text{ef}\,u}\right]$ *gleich* 1 *wird.*

Bestimmen wir nämlich die reellen Zahlen x, y so, daß

$$\omega_f - \omega_e = 2x\,\omega_c + 2y\,\omega_n, \qquad \text{also} \qquad B = \frac{\pi i}{\omega_c}\,y \tag{5.3}$$

wird, so sieht man, daß die Funktionalgleichungen

$$\text{ef}(u+2\omega_c) = e^{2\pi i y}\,\text{ef}\,u, \tag{5.4}$$

$$\text{ef}(u+2\omega_n) = e^{-2\pi i x}\,\text{ef}\,u \tag{5.5}$$

gelten. Es ist umgekehrt nicht schwer, in (5.1) die Konstante B so zu bestimmen, daß ef u doppeltperiodischen absoluten Betrag hat, und dabei ergibt sich zwangsläufig (5.2).

Wenn die Funktion ef u für verschiedene Werte von ω_e, ω_f betrachtet werden soll, so sollen an Stelle der Buchstaben e, f auch andere treten können, z.B. g, h, j. Die Buchstaben s, n, c, d sollen jedoch nur verwendet werden, wenn $\omega_s, \omega_n, \omega_c, \omega_d$ die bisher gültige Bedeutung hat.

Sind e, f *zwei dieser Buchstaben* s, c, n, d, *so erhalten wir genau die zwölf Jacobi-Glaisherschen Funktionen.*

Allgemein gelten die Regeln

$$\text{ef}\,u\,\text{fg}\,u = \text{eg}\,u, \qquad \frac{1}{\text{ef}\,u} = \text{fe}\,u, \qquad \text{ef}\,u\,\text{gh}\,u = \text{eh}\,u\,\text{gf}\,u. \tag{5.6}$$

2. Anwendung auf die Transformation der elliptischen Funktionen. Die in (5.3) eingeführten Zahlen x, y seien jetzt rational, etwa

$$x = \frac{a}{m}, \qquad y = \frac{b}{m}, \tag{5.7}$$

und der größte gemeinsame Teiler (a, b, m) von a, b und m sei 1. Weil sich $\mathrm{ef}\, u$ bei Vermehrung von x und y um ganze Zahlen nicht ändert, ist es keine Beschränkung der Allgemeinheit, a und b als teilerfremd anzunehmen [andernfalls setze man $b = b'b''$, $b' \mid m$, $(b'', m) = 1$ und ersetze a durch eine Lösung a_1 der Kongruenzen $a_1 \equiv a \pmod m$, $a_1 \equiv 1 \pmod{b''}$. Dann ist $(a_1, b) = 1$]. Für ganze Zahlen M, N ist

$$\mathrm{ef}(u + 2M\omega_c + 2N\omega_n) = e^{\frac{2\pi i}{m}(Mb - Na)}\, \mathrm{ef}\, u. \tag{5.8}$$

$\mathrm{ef}\, u$ ist also doppeltperiodisch zu einem Teilgitter $\dot{\mathfrak{W}}$ des Gitters $\mathfrak{W}$, das seinerseits das von $2m\omega_c$ und $2m\omega_n$ erzeugte Gitter $m\mathfrak{W}$ enthält. $Mb - Na$ durchläuft wegen $(a, b, m) = 1$ alle Reste mod m. Diese Reste entsprechen umkehrbar eindeutig den Nebenklassen von $\dot{\mathfrak{W}}$ in $\mathfrak{W}$, *d.h.* $\dot{\mathfrak{W}}$ *ist ein Teilgitter von* $\mathfrak{W}$ *vom Index* m, *und* $m\mathfrak{W}$ *ist seinerseits Teilgitter von* $\dot{\mathfrak{W}}$ *vom Index* m.

Offenbar erhält man für geeignete a, b jedes Teilgitter dieser Eigenschaften, das in keinem Gitter $k\mathfrak{W}$ mit $k > 1$ liegt. Es gibt primitive Perioden

$$2\dot{\omega}_c = 2a\omega_c + 2b\omega_n, \qquad 2\dot{\omega}_n \in m\mathfrak{W}$$

von $\mathrm{ef}\, u$. Wegen $(a, b) = 1$ ist $2h\dot{\omega}_c \in m\mathfrak{W}$, wenn und nur wenn $h \equiv 0 \pmod m$ ist. Es sei $\frac{2\dot{\omega}_n}{m} = 2\omega_0 \in \mathfrak{W}$. Dann ist

$$\mathrm{ef}(u + 2\dot{\omega}_c) = \mathrm{ef}\, u, \qquad \mathrm{ef}(u + 2\omega_0) = \zeta \cdot \mathrm{ef}\, u, \tag{5.9}$$

wobei ζ eine primitive m-te Einheitswurzel ist. $2\omega_0$ und $2\dot{\omega}_c$ erzeugen $\mathfrak{W}$.

$\mathsf{E}_{\mathfrak{W}}$, $\mathsf{E}_{\dot{\mathfrak{W}}}$ *seien die Körper der elliptischen Funktionen zu den Periodengittern* $\mathfrak{W}$ *bzw.* $\dot{\mathfrak{W}}$. $\mathsf{E}_{\dot{\mathfrak{W}}}$ *ist Erweiterung* m*-ten Grades von* $\mathsf{E}_{\mathfrak{W}}$. Denn jedes Element $f(u) \in \mathsf{E}_{\dot{\mathfrak{W}}}$ genügt einer Gleichung m-ten Grades mit Koeffizienten aus $\mathsf{E}_{\mathfrak{W}}$, deren Wurzeln $f(u), f(u + 2\omega_0), \ldots, f(u + 2(m-1)\omega_0)$ sind. Folglich ist

$$\mathsf{E}_{\dot{\mathfrak{W}}} = \mathsf{E}_{\mathfrak{W}}(\mathrm{ef}\, u).$$

Denn $\mathrm{ef}\, u$ genügt der Gleichung

$$x^m = (\mathrm{ef}\, u)^m \in \mathsf{E}_{\mathfrak{W}}.$$

Sie ist irreduzibel, denn ihre Wurzeln sind $\mathrm{ef}\, u \cdot \zeta^h$ $(h = 0, 1, \ldots, m-1)$, und ein irreduzibler Faktor k-ten Grades hätte das konstante Glied $\zeta^r (\mathrm{ef}\, u)^k$ [mit ganzzahligem r], das für $k < m$ nicht in $\mathsf{E}_{\mathfrak{W}}$ liegt. *Mit Hilfe der Funktion* ef u *kann diese Körpererweiterung leicht rechnerisch durchgeführt werden.* Es sei nämlich $f(u) \in \mathsf{E}_{\dot{\mathfrak{W}}}$. Wir betrachten die Funktionen $(h = 0, 1, \ldots, m-1)$

$$f_h(u) = f(u) + \zeta^h f(u + 2\omega_0) + \zeta^{2h} f(u + 4\omega_0) + \cdots + \zeta^{(m-1)h} f[u + 2(m-1)\omega_0]. \tag{5.10}$$

Dann ist

$$f_h(u + 2\dot{\omega}_c) = f_h(u), \quad f_h(u + 2\omega_0) = \zeta^{-h} \cdot f_h(u).$$

$g_h(u) = f_h(u)\,(\mathrm{ef}\,u)^h$ sind nach (5.9) Funktionen aus $\mathsf{E}_{\mathfrak{M}}$, deren Pole mit ihren Hauptteilen durch rationale Rechnungen bestimmt werden können und die daher z. B. in der Form $\sum_i a_i\,\mathrm{zs}(u - b_i) + \sum_i a_i'\,\mathrm{zs}'(u - b_i) + \cdots$ darstellbar sind. Dann wird

$$f(u) = \frac{1}{m} \sum_{h=0}^{m-1} g_h(u)\,(\mathrm{ef}\,u)^h. \tag{5.11}$$

Dabei wurde als Konstantenkörper der Körper der komplexen Zahlen genommen. Die Werte $\mathrm{ef}(2h\omega_0)$, $h = 0, 1, \ldots, m-1$ gehen in die Rechnung ein. Die mit der algebraischen Bestimmung dieser Werte verbundenen Schwierigkeiten können natürlich so wenig wie bei den Jacobischen oder Weierstraßschen Funktionen zum Gitter $\dot{\mathfrak{M}}$ durch die einfachen Mittel der vorliegenden Arbeit überwunden werden, sondern führen auf die Theorie der Modulfunktionen und fallen damit aus dem Rahmen dieser Arbeit, die eine *vereinfachte Darstellung der Elemente der Theorie* geben wollte[11]).

[11]) Über die vorliegende Arbeit wurde Anfang 1956 in einem Vortrag im Münchener mathematischen Kolloquium und am 18. 9. 56 in einem Kurzvortrag auf dem IV. Österreichischen Mathematikerkongreß in Wien berichtet.

München 19, Pötschnerstr. 8

(Eingegangen am 13. April 1956)

PERRON, O.
Math. Zeitschr. Bd. 67, S. 176—180 (1957)

Ein neuartiges diophantisches Problem

Dem jugendlichen GEORG FABER zum 80. Geburtstag am 5. April 1957

Von
OSKAR PERRON

I. Sind a, b zwei verschiedene reelle Zahlen und durchläuft in der Funktion

$$f(x) = |a - x| \cdot |b - x|$$

x alle ganzen rationalen Zahlen, so ist klar, daß es ein Minimum gibt. Denn $f(x)$ wird für hinreichend große $|x|$ beliebig groß, so daß man nur endlich viele Zahlen x durchzuprobieren braucht, um bestimmt auf das Minimum von $f(x)$ zu stoßen. Das ist eine triviale Sache. Der hypertriviale Fall, daß eine der Zahlen a, b selbst rational ganz, das Minimum also gleich 0 ist, mag im folgenden ausgeschlossen bleiben. Da eine der Zahlen a, b die kleinere sein muß und da sich die Frage nicht ändert, wenn man beide um die gleiche ganze Zahl vermehrt, kann man $0 < a < 1$, $a < b$ voraussetzen; das soll im folgenden geschehen. Ich bin nun im Zusammenhang mit einer ganz anderen Aufgabe auf die Frage geführt worden, ob das Minimum ≤ 1 ist. Das ist trivialerweise der Fall, wenn die Differenz $b - a$ klein ist, etwa kleiner als $\frac{1}{2}$. Dann macht nämlich die der Zahl a nächstgelegene ganze Zahl x jeden der beiden Faktoren von $f(x)$ kleiner als 1. Auch wenn $b - a = c$ sehr groß ist, kann es sein, daß das Minimum kleiner als 1 ist, sobald nämlich eine der Zahlen a, b hinreichend nahe bei einer ganzen Zahl liegt; im allgemeinen wird es aber nicht so sein. Das ist wieder trivial. Man kann nun die Frage stellen, ob bei *jedem* Zahlenpaar a, b mit *vorgegebener* Differenz c das Minimum ≤ 1 ist. Damit hört aber die Trivialität auf. Der folgende Satz beantwortet die Frage teilweise:

Satz 1. *Wenn* $0 < a < 1$, $0 < c \leq \sqrt{5}$, *dann ist das Minimum von*

$$f(x) = |a - x| \cdot |a + c - x|$$

kleiner als 1, *wenn man von dem Fall* $a = \frac{3 - \sqrt{5}}{2}$, $c = \sqrt{5}$ *absieht; in diesem ist das Minimum gleich* 1.

Beweis. Ist g die größte in $a + c$ enthaltene ganze Zahl, also $g < a + c < g + 1$, so ist klar, daß das Minimum, falls es ≤ 1 ist, nur von einer (oder mehreren) der vier Zahlen $0, 1, g, g + 1$ geliefert werden kann. Denn für alle anderen x sind beide Faktoren von $f(x)$ größer als 1. Offenbar ist $g \geq 0$. Für $g = 0$ ist schon $f(0) = a(a + c) < 1 \cdot 1$, und es ist nichts mehr zu beweisen. Sei daher $g \geq 1$.

Nun ist, wenn wir von der Voraussetzung $c \leq \sqrt{5}$ vorläufig ganz absehen,

$$(1)\qquad \begin{cases} f(0) &= a(a+c), \\ f(1) &= (1-a)(a+c-1), \\ f(g) &= (g-a)(a+c-g), \\ f(g+1) &= (g+1-a)(g+1-a-c). \end{cases}$$

Wenn wir nun einmal annehmen, das Minimum sei ≥ 1, so ist $1 \leq f(x)$ für alle x, also insbesondere für $x = 0, 1, g, g+1$. Daher ist

erstens: $$1 \leq f(0) = a(a+c) = \left(a + \frac{c}{2}\right)^2 - \frac{c^2}{4},$$

und daraus folgt

$$(2)\qquad a + \frac{c}{2} \geq \sqrt{\frac{c^2}{4} + 1},$$

zweitens: $$1 \leq f(1) = (1-a)(a-1+c) = \frac{c^2}{4} - \left(a - 1 + \frac{c}{2}\right)^2,$$

und daraus folgt $\frac{c^2}{4} \geq 1$ und

$$(3)\qquad a - 1 + \frac{c}{2} \leq \sqrt{\frac{c^2}{4} - 1},$$

drittens: $$1 \leq f(g) = (g-a)(a-g+c) = \frac{c^2}{4} - \left(g - a - \frac{c}{2}\right)^2,$$

und daraus folgt $\frac{c^2}{4} \geq 1$ und

$$(4)\qquad g - a - \frac{c}{2} \leq \sqrt{\frac{c^2}{4} - 1},$$

viertens: $1 \leq f(g+1) = (g+1-a)(g+1-a-c) = \left(g + 1 - a - \frac{c}{2}\right)^2 - \frac{c^2}{4}$,

und daraus folgt

$$(5)\qquad g + 1 - a - \frac{c}{2} \geq \sqrt{\frac{c^2}{4} + 1}.$$

Aus (2) und (3) oder auch aus (5) und (4) ergibt sich

$$\sqrt{\frac{c^2}{4} + 1} \leq 1 + \sqrt{\frac{c^2}{4} - 1},$$

also, wenn man die rechtsstehende Wurzel nach links bringt und dann die Ungleichung mit der Summe der beiden Wurzeln multipliziert:

$$2 \leq \sqrt{\frac{c^2}{4} + 1} + \sqrt{\frac{c^2}{4} - 1}.$$

Daraus folgt, da die rechtsstehende Funktion von c mit c wächst und für $c = \sqrt{5}$ den Wert 2 hat, sofort: $c \geq \sqrt{5}$. Für $c < \sqrt{5}$ kann also die oben gemachte

Annahme, das Minimum sei ≥ 1, nicht zutreffen, womit der Hauptteil von Satz 1 bereits bewiesen ist. Für $c = \sqrt{5}$ folgt aus (2) und (3)

$$a \geq \frac{3-\sqrt{5}}{2}, \quad a \leq \frac{3-\sqrt{5}}{2}.$$

Also kann für $c = \sqrt{5}$ die Annahme, das Minimum sei ≥ 1, ebenfalls nicht zutreffen, wenn man von der Zahl $a = \frac{3-\sqrt{5}}{2}$ absieht. Für diese ist dann $a + c = \frac{3+\sqrt{5}}{2}$, also $g = 2$, und alle vier Ausdrücke (1) sind gleich 1, womit auch der Rest von Satz 1 bewiesen ist.

II. Kenner diophantischer Approximationen werden durch das hier beobachtete Auftreten der Wurzel aus 5 gewiß daran erinnert, daß dieselbe Zahl auch in dem bekannten Satz von HURWITZ auftritt: Jede Irrationalzahl ξ läßt sich durch unendlich viele rationale Brüche $\frac{p}{q}$ so approximieren, daß der Fehler kleiner wird als $\frac{1}{\sqrt{5}\,q^2}$, und hier läßt sich die Zahl $\sqrt{5}$ durch keine für alle ξ gültige größere ersetzen (HURWITZ [3], Beweise unter anderen auch bei FUJIWARA [2], PERRON [5], [7]). Das ist nun trotz der Verschiedenartigkeit der Probleme kein neckischer Zufall; vielmehr bin ich gerade durch Beschäftigung mit dem HURWITZschen Satz zu meiner Fragestellung gelangt und habe das Resultat vorausgesehen.

Der HURWITZsche Satz hat ein Analogon in jedem imaginären quadratischen Zahlkörper, doch kennt man die der $\sqrt{5}$ entsprechende Schranke nur in wenigen Fällen. Speziell im Körper $\mathfrak{R}(i)$ ist sie von FORD [1] und auf anderem Weg auch von mir [4] gefunden worden, und zwar lautet der betreffende Satz so: Jede komplexe Zahl ξ, die nicht dem Körper $\mathfrak{R}(i)$ angehört, läßt sich durch unendlich viele Brüche $\frac{p}{q}$, deren Zähler und Nenner ganze Zahlen aus $\mathfrak{R}(i)$ sind, so approximieren, daß der Fehler kleiner als $\frac{1}{\sqrt{3}\,|q|^2}$ wird, und hier läßt sich die Zahl $\sqrt{3}$ durch keine für alle ξ gültige größere Zahl ersetzen. Auch mit dieser Wurzel aus 3 kann Wiedersehen feiern, wer sich die folgende Frage stellt: Wie groß darf die Zahl c sein, damit bei jedem Paar komplexer Zahlen a, b, für das $|b-a| = c$ ist, stets die Ungleichung $|a-x| \cdot |b-x| \leq 1$ durch eine ganze Zahl x aus $\mathfrak{R}(i)$ befriedigt werden kann? Die Antwort lautet nämlich: *Für $c \leq \sqrt{3}$ ist das stets möglich, und zwar mit dem Kleinerzeichen, wenn man von den Fällen*

(I) $$a \equiv \frac{i \pm \sqrt{3}}{2}, \quad b = a \mp \sqrt{3},$$

(II) $$a \equiv \frac{1 \pm i\sqrt{3}}{2}, \quad b = a \mp i\sqrt{3}$$

absieht; für diese geht es bloß mit dem Gleichheitszeichen (das Kongruenzzeichen $\equiv d$ bedeutet hier natürlich soviel wie $= d +$ eine ganze Zahl aus $\mathfrak{R}(i)$).

Von dem nicht ganz einfachen Beweis dieses Satzes will ich hier absehen. Es ist übrigens nicht so, daß auch bei den anderen imaginären quadratischen Zahlkörpern die entsprechenden Schranken für die beiden Probleme einander gleich sind. Im Körper $\mathfrak{R}(i\sqrt{3})$, wo die der HURWITZschen $\sqrt{5}$ entsprechende Schranke gleich $\sqrt[4]{13}$ ist (PERRON [5]), mag es noch so sein. Aber schon bei $\mathfrak{R}(i\sqrt{2})$, wo die Schranke gleich $\sqrt{2}$ ist (PERRON [6]), glaube ich es nicht, ebensowenig bei $\mathfrak{R}(i\sqrt{7})$ und $\mathfrak{R}(i\sqrt{11})$; und bei den anderen Körpern ist das Problem in dieser Form überhaupt gegenstandslos, weil das Gitter der ganzen Zahlen so weitmaschig ist, daß man beliebig nahe beieinander gelegene Zahlen a, b angeben kann, die von allen ganzen x einen Abstand größer als 1 haben. Trotzdem dürfte für alle Körper eine Übereinstimmung gelten, wenn man die Funktion $f(x)$ in gewisser Weise modifiziert. Beim Körper $\mathfrak{R}(i\sqrt{2})$ habe ich z.B. Ursache zu der Vermutung, daß Übereinstimmung stattfindet, wenn man an Stelle der Funktion $f(x)$ die folgende betrachtet:

$$\operatorname{Min}\left\{f(x),\ 2f\left(x+\frac{i}{\sqrt{2}}\right),\quad 2f\left(x-\frac{i}{\sqrt{2}}\right)\right\}.$$

III. Es gibt nun aber noch eine andere Überraschung. Nachdem durch Satz 1 festgestellt ist, daß die Ungleichung $|a-x|\cdot|b-x|<1$ für $|b-a|<\sqrt{5}$ immer erfüllbar ist, für $|b-a|=\sqrt{5}$ dagegen nicht mehr immer, wird man es gerne für selbstverständlich halten, daß für $|b-a|>\sqrt{5}$ die Ungleichung erst recht nicht mehr immer erfüllbar ist. *Aber das ist gerade falsch*, wie jetzt noch gezeigt werden soll.

Zu dem Zweck nehmen wir den Beweis von Satz 1 wieder auf, gehen also von der Annahme aus, das Minimum sei ≥ 1. Dann gelten die Formeln (2) bis (5). Aus (5) und (2) folgt

$$\sqrt{\frac{c^2}{4}+1}\leq g+1-a-\frac{c}{2}\leq g+1-\sqrt{\frac{c^2}{4}+1},$$

also ist $2\sqrt{\frac{c^2}{4}+1}\leq g+1$, und folglich

$$\frac{c^2}{4}\leq\left(\frac{g+1}{2}\right)^2-1. \tag{6}$$

Aus (3) und (4) folgt

$$\sqrt{\frac{c^2}{4}-1}\geq a+\frac{c}{2}-1\geq g-1-\sqrt{\frac{c^2}{4}-1},$$

also ist $2\sqrt{\frac{c^2}{4}-1}\geq g-1$, und folglich

$$\frac{c^2}{4}\geq\left(\frac{g-1}{2}\right)^2+1. \tag{7}$$

Aus (6) und (7) folgt nun

$$\left(\frac{g+1}{2}\right)^2-1\geq\left(\frac{g-1}{2}\right)^2+1,$$

oder einfacher geschrieben:

$$g\geq 2.$$

Wenn nun $g=2$ ist, so folgt aus (6) und (7): $c=\sqrt{5}$. Das ist der Fall, den wir schon kennen. Wenn aber $g>2$ ist, so ist g mindestens gleich 3, und aus (7) folgt dann: $c\geq\sqrt{8}$. Wenn daher $c>\sqrt{5}$ ist, so muß, wenn die Annahme, das Minimum sei ≥ 1, zutreffen soll, sogar $c\geq\sqrt{8}$ sein. Für $\sqrt{5}<c\leq\sqrt{8}$ ist also das Minimum stets kleiner als 1, mit allenfallsiger Ausnahme des Falles $c=\sqrt{8}$, in dem dann aus (3) und (4) folgt:

$$a\leq 2-\sqrt{2},\quad a\geq g-1-\sqrt{2}.$$

Daher muß $g=3$ und $a=2-\sqrt{2}$ sein. In diesem Fall sind $f(0)$ und $f(g+1)$ gleich 2, aber $f(1)$ und $f(g)$ gleich 1, das Minimum ist also gleich 1. Damit ist bewiesen der

Satz 2. *Wenn* $0<a<1$, $\sqrt{5}<c\leq\sqrt{8}$, *so ist stets das Minimum von* $|a-x|\cdot|a+c-x|$ *kleiner als* 1, *wenn man von dem Fall* $a=2-\sqrt{2}$, $c=\sqrt{8}$ *absieht; in diesem ist das Minimum gleich* 1.

Auch in diesem merkwürdigen Sprung von $\sqrt{5}$ auf $\sqrt{8}$ wird einen alten Bekannten begrüßen, wer sich an den weiteren Satz von HURWITZ erinnert: Diejenigen Irrationalzahlen, die sich, wenn d eine Zahl größer als $\sqrt{5}$ ist, durch unendlich viele rationale Brüche so gut approximieren lassen, daß der Fehler kleiner als $\frac{1}{dq^2}$ ist, gestatten sogar unendlich viele Approximationen mit einem Fehler kleiner als $\frac{1}{\sqrt{8}\,q^2}$ (HURWITZ [3], Beweise unter anderen auch bei FUJIWARA [2], PERRON [7]).

Literatur

[1] FORD, L. R.: On the closeness of approach of complex rational fractions to a complex irrational number. Trans. Amer. Math. Soc. **27** (1927). — [2] FUJIWARA, M.: Bemerkung zur Theorie der Approximation der irrationalen Zahlen durch rationale Zahlen. Tohoku Math. J. **11** (1917); **14** (1918). — [3] HURWITZ, A.: Über die angenäherte Darstellung der Irrationalzahlen durch rationale Brüche. Math. Ann. **39** (1891) = Werke II. — [4] PERRON, O.: Über die Approximation einer komplexen Zahl durch Zahlen des Körpers $\mathfrak{R}(i)$. Math. Ann. **103** (1930); **105** (1931). — [5] PERRON, O.: Über einen Approximationssatz von Hurwitz und über die Approximation einer komplexen Zahl durch Zahlen des Körpers der dritten Einheitswurzeln. Sitzgsber. Bayer. Akad. Wiss., Math.-nat. Abt. **1931**. — [6] PERRON, O.: Diophantische Approximationen in imaginären quadratischen Zahlkörpern, insbesondere im Körper $\mathfrak{R}(i\sqrt{2})$. Math. Z. **37** (1933). — [7] PERRON, O.: Die Lehre von den Kettenbrüchen, 3. Aufl. Bd. 1. Stuttgart 1954.

München 27, Friedrich-Herschel-Str. 11

(Eingegangen am 22. April 1956)

STEUERWALD, R.
Math. Zeitschr. Bd. 67, S. 181—187 (1957)

Über die Perioden regelmäßiger Kettenbrüche für Quadratwurzeln aus ganzen Zahlen. II

Herrn GEORG FABER zum 80. Geburtstag am 5. April 1957 gewidmet

Von

RUDOLF STEUERWALD

Einleitung

Diese Arbeit soll eine früher unter dem gleichen Titel erschienene [Math. Z. **52**, 686—697 (1950)] ergänzen und ist mit dieser so eng verzahnt, daß dem Leser der Rückgriff darauf nicht erspart werden kann. Deshalb werden auch hier die Formeln usw. nicht wieder von vorn, sondern anschließend an die früheren Zahlenwerte numeriert. Es sei daher ein für allemal bemerkt: Hinweise auf die Paragraphen 1 bis 6, die Sätze 1 bis 6 und 1* bis 5*, die Definitionen 1 und 2, die Formeln (1) bis (28) beziehen sich auf die ältere, Hinweise mit höheren Nummern auf die vorliegende Arbeit. Nur bei Seitenangaben wird erstere durch I, letztere durch II gekennzeichnet.

Hier soll gezeigt werden, wie man die früher nur über „ausgezeichnete" Q, c-Folgen gemachten Aussagen (vgl. I, 691 ff., § 4: Def. 1 und 2, sowie Abschnitt III) teilweise auf allgemeine Q, c-Folgen (§ 7) und vollständig auf eine andere besondere Art von solchen (§ 8) sinngemäß ausdehnen kann. Diese Ergebnisse gestatten dann, den Fall $Q=3$ entsprechend zu erledigen (§ 9) wie seinerzeit den Fall $Q=2$.

§ 7

In diesem Paragraphen wird das Analogon zu Satz 4 (I 692) für allgemeine Q, c-Folgen r-ter Ordnung ($r \geqq 3$) entwickelt. Dabei werden folgende Abkürzungen benützt:

$$(29) \qquad \begin{cases} q = Q - c\,b_{r-1}, & U = q\,b_1 - b_{r-1}, & V = q\,b_{r-2} - c, \\ W = b_{r-1}b_{r-2} - q\,b_1 b_{r-2} + c\,b_1 + 1. \end{cases}$$

Zwischen den Größen U, V, W bestehen die Beziehungen

$$(30) \qquad W = -U\,b_{r-2} + c\,b_1 + 1 = -V\,b_1 + b_{r-1}b_{r-2} + 1,$$

$$(31) \qquad q\,W = Q - U\,V.$$

Wir werden nun eine Reihe von Ungleichungen aufstellen und beweisen, aber mit Mitteln, die teilweise nur für $r \geqq 4$ funktionieren und für $r=3$ versagen. Die Einbeziehung dieses letzteren Falles durch direkte Verifikation darf aber billigerweise dem Leser überlassen werden; denn sie ist ganz einfach wenn man beachtet, daß aus (18), (19) (s. I 690) und (30) folgt:

Für $r=3$ ist $U=c$, $V=b_2$, $W=1$.

Aus (18) schließt man mittels (14) und $q \geq 1$ (s. Satz 1, I 691) auf

$$(-U + c b_1 + 1) B_{r-3,1} + (c - q) B_{r-5,2} > 0 \quad (r \geq 4)$$

und daraus, falls $q - c \geq 0$, auf $U \leq c b_1$ oder

$$b_{r-1} \geq (q - c) b_1. \tag{32}$$

Diese Ungleichung gilt aber trivialerweise auch, falls $q - c < 0$.

Ähnlich findet man mittels (19) $V \leq b_{r-1} b_{r-2}$ oder

$$c \geq (q - b_{r-1}) b_{r-2}. \tag{33}$$

Ferner folgt, wieder aus (18) mittels (14)

$$-(U + q) B_{r-3,1} + (c b_1 + c + 1) B_{r-5,2} < 0 \quad (r \geq 4)$$

und daraus

$$U \geq 1 - q; \tag{34}$$

analog führt (19) auf

$$V \geq 1 - q. \tag{35}$$

Endlich gilt

$$W \geq 1. \tag{36}$$

Diese Beziehung folgt nämlich, wenn wenigstens eine der Ungleichungen $U \leq 0$, $V \leq 0$ besteht, aus (30); wenn aber $U > 0$ und $V > 0$ ist, aus (20) (I 690).

Jetzt führen (36), (16) und (31) auf

$$UV \leq c b_{r-1}. \tag{37}$$

Nunmehr beweisen wir:

$$W - U - V - q \leq -1. \tag{38}$$

Ist zunächst $r \geq 5$, so schließt man aus (20) mittels (14) auf

$$(W - V) B_{r-5,2} - (U + q) B_{r-5,1} < 0, \tag{39}$$

$$(W - U) B_{r-5,1} - (V + q) B_{r-5,2} < 0. \tag{40}$$

Je nachdem nun $B_{r-5,1} \leq B_{r-5,2}$ oder $B_{r-5,1} \geq B_{r-5,2}$ ist, folgt (38) aus (39) mittels (34) oder aus (40) mittels (35).

Ist aber $r = 4$, so führt (20) auf $W = q$; und damit verifiziert man

$$(U + V)(b_1 b_2 + 1) = b_1 + b_2 + b_3 (b_2^2 - 1) + c (b_1^2 - 1) \geq 2,$$

also auch $U + V \geq 1$.

Durch Zusammenfassung der Beziehungen (32) bis (35) und (37) sowie der beiderseits mit q multiplizierten und dann mittels (31) umgeformten Ungleichung (38) erhält man jetzt:

Satz 7. *In jeder Q, c-Folge r-ter Ordnung ($r \geq 3$) ist, falls man $Q - c b_{r-1} = q$ setzt,*

$$c b_1 \geq q b_1 - b_{r-1} \geq 1 - q, \quad b_{r-1} b_{r-2} \geq q b_{r-2} - c \geq 1 - q,$$
$$(q b_1 - b_{r-1})(q b_{r-2} - c) \leq Q - q < Q + q \leq (q b_1 - b_{r-1} + q)(q b_{r-2} - c + q).$$

§ 8

Früher (I 691: Def. 2) wurden als „ausgezeichnete" Q, c-Folgen diejenigen bezeichnet, bei denen $c > 0$ ist und b_{r-1} seine *obere* Schranke $\frac{Q-1}{c}$ erreicht. Nur für sie konnten in Satz 5 (I 693) gewisse Rekursionsbeziehungen nachgewiesen werden. Hier soll nun gezeigt werden, daß entsprechende Aussagen auch für diejenigen Q, c-Folgen mit $c > 0$ gelten, bei denen b_{r-1} seine *untere* Schranke $\frac{Q-c}{c+1}$ (vgl. I 691, Satz 1) erreicht. Um auch für sie einen bequemen Namen zu haben, sei es erlaubt, die früheren „ausgezeichneten" Q, c-Folgen von jetzt ab statt dessen „*maximale*" und die neuen „*minimale*" zu nennen.

Definition 3. *Eine Q, c-Folge r-ter Ordnung, bei der $c > 0$, $c+1$ Teiler von $Q - c$ und $b_{r-1} = \frac{Q-c}{c+1}$ ist, heiße eine minimale Q, c-Folge.*

Bei einer minimalen Q, c-Folge wird gemäß ihrer Definition

$$q = c + b_{r-1}, \quad Q + 1 = (c+1)(b_{r-1} + 1). \tag{41}$$

Damit folgt aus (32) und (33) für $r \geq 3$:

$$b_1 = b_{r-2} = 1, \tag{42}$$

also

$$U = c, \quad V = b_{r-1}, \quad W = 1. \tag{43}$$

Jetzt setzen wir in (17) (I 690) $q = c + b_{r-1}$, wenden dann (12) auf $c B_{r-2}$, (13) auf $b_{r-1} B_{r-2}$ an, berücksichtigen (42) und (43) und multiplizieren beiderseits mit -1. Dann ergibt sich

$$c B_{r-4} + b_{r-1} B_{r-4,2} - B_{r-4,1} = 0 \quad (r \geq 3). \tag{44}$$

Daraus erhält man: erstens, indem man (12) auf $B_{r-4,2}$, (13) auf $B_{r-4,1}$, B_{r-4} und $B_{r-5,1}$ anwendet,

$$(b_{r-1} - b_2) B_{r-5,2} + (c b_2 + q) B_{r-6,2} + c B_{r-7,3} - B_{r-6,3} = 0 \quad (r \geq 5); \tag{45}$$

zweitens, indem man (13) auf B_{r-4}, (12) auf $B_{r-4,1}$, $B_{r-4,2}$ und $B_{r-5,2}$ anwendet,

$$(c - b_{r-3}) B_{r-5,1} + (b_{r-1} b_{r-3} + q) B_{r-6,2} + b_{r-1} B_{r-7,2} - B_{r-6,1} = 0 \quad (r \geq 5). \tag{46}$$

Endlich geht (45) durch Anwendung von (12) auf $B_{r-5,2}$ und $B_{r-6,3}$ nach leichter Umformung über in

$$\left\{ \begin{aligned} &[Q - (b_{r-1} - b_2)(c - b_{r-3})] B_{r-6,2} + (b_{r-1} - b_2) B_{r-7,2} \\ &\qquad + (c - b_{r-3}) B_{r-7,3} - B_{r-6,3} = 0 \end{aligned} \right. \quad (r \geq 6). \tag{47}$$

Die Analogie zwischen (24) bis (26) (I 693) und (45) bis (47) ist evident; daher führt auch eine Schlußweise, welche der dortigen genau entspricht, in Verbindung mit (42) zu den beiden folgenden Sätzen.

Satz 8. *In jeder minimalen Q, c-Folge r-ter Ordnung ($r \geq 5$) ist*

$$b_1 = b_{r-2} = 1, \qquad b_2 - b_{r-1} \geq 0, \qquad b_{r-3} - c \geq 0,$$

$$(b_2 - b_{r-1})(b_{r-3} - c) < Q < (b_2 - b_{r-1} + 1)(b_{r-3} - c + 1).$$

Insbesondere gilt also: Wenn von den beiden nicht negativen Differenzen $b_2 - b_{r-1}$, $b_{r-3} - c$ die eine nicht verschwindet, ist die andere $\leq Q - 1$; wenn die eine verschwindet, ist die andere $\geq Q$.

Satz 9. *Eine Folge $b_1, \ldots, b_{r-1} = \frac{Q-c}{c+1}$ ($c > 0$, $r \geq 7$) ist dann und nur dann eine minimale Q, c-Folge, wenn $b_1 = b_{r-2} = 1$,*

ferner $b_3, \ldots, b_{r-4}$, c^* eine Q, c'-Folge

oder $b_{r-4}, \ldots, b_3$, c' eine Q, c^*-Folge

und $b_2 = b_{r-1} + c'$, $b_{r-3} = c + c^*$ ist.

Zusatz. Der Satz gilt auch noch für $r = 6$, wenn man $b_3, \ldots, b_{r-4}$ als leere Folge deutet, und liefert dann die Bedingung $c' c^* = Q - 1$.

Anmerkung. Von den beiden Bedingungen in Zeile 3 und 4 des Satzes 9 kommt, falls $c' = 0$, nur die erste; falls $c^* = 0$, nur die zweite in Frage. Ist $c' > 0$ und $c^* > 0$, so sind sie nach Satz 2 (I 692) äquivalent; $c' = c^* = 0$ ist durch Satz 8 ausgeschlossen.

Erwähnt sei noch der unmittelbar evidente

Satz 10. *Ist von den beiden Folgen des Satzes 2* (I 692) *die eine maximal oder minimal, so hat die andere die nämliche Eigenschaft.*

Dabei ist zu beachten, daß aus den Definitionen folgt: Keine Q, c-Folge kann gleichzeitig maximal und minimal sein.

Aus den Beispielen des § 2 (I 688—689) ersieht man:

Für $r = 1$ und $r = 2$ existiert keine minimale Q, c-Folge, für $r = 3$ höchstens eine mit

$$b_1 = 1, \qquad b_2 = \frac{Q-c}{c+1}.$$

Weiter findet man unter Benützung des Satzes 8 durch leichte Rechnung: Auch für $r = 5$ existiert höchstens eine minimale Folge mit

$$b_1 = b_3 = 1, \qquad b_2 - c = b_4 = \frac{Q-c}{c+1}.$$

Daher ist es bemerkenswert, daß für $r = 4$ keine minimale Q, c-Folge existiert. Das sieht man so: Die entsprechende Gleichung (I 688) geht, wenn man in ihr

$$b_3 = \frac{Q-c}{c+1}$$

setzt und den Nenner wegschafft, über in

$$(48) \qquad Q(b_1b_2 - b_2 + 1) + c^2(b_1b_2 - b_1 + 1) + c(b_2 - b_1 - 1) - 1 = 0.$$

Für jedes Paar natürlicher Zahlen b_1, b_2 wird

$$b_1b_2 - b_2 + 1 \geq 1, \quad b_1b_2 - b_1 + 1 \geq 1,$$

also, wenn man noch beachtet, daß $1 \leq c \leq Q-1$ ist, die linke Seite von (48) sicher größer als

$$c(b_1b_2 - b_2 + 1) + c(b_1b_2 - b_1 + 1) + c(b_2 - b_1 - 1) - c = 2cb_1(b_2 - 1) \geq 0.$$

Somit hat die Gl. (48) keine Lösung in unserem Sinn.

§ 9

Die vorstehende Ergänzung der früheren Ergebnisse erlaubt jetzt, auch für $Q = 3$ die Q, c-Folgen in Klassen einzuteilen, da hier nur folgende vier Fälle möglich sind:

1. $c = 0$, 2. $c = 1$ minimal, 3. $c = 1$ maximal, 4. $c = 2$ maximal.

Für die entsprechend numerierten Klassenzahlen $n_j(r)$ $(j = 1, 2, 3, 4)$ liefern die Sätze 3, 5 und 9 die Rekursionsformeln

$$(49) \quad n_1(r) = n_1(r-2) + n_2(r-2) + n_3(r-2) + n_4(r-2) \qquad (r \geq 4),$$

$$(50) \quad n_2(r) = 2n_1(r-4) + n_2(r-4) + n_3(r-4) + n_4(r-4) \qquad (r \geq 6),$$

$$(51) \quad n_3(r) = n_4(r) = 2n_1(r-2) + n_2(r-2) + n_3(r-2) + n_4(r-2) \qquad (r \geq 4).$$

Subtrahiert man einerseits (49) von (51), andererseits (49) mit $r-2$ statt r von (50), so erhält man

$$(52) \qquad n_3(r) = n_4(r) = n_1(r) + n_1(r-2) \qquad (r \geq 4),$$

$$(53) \qquad n_2(r) = n_1(r-2) + n_1(r-4) \qquad (r \geq 6).$$

Schreibt man jetzt (52) und (53) mit $r-2$ statt r und setzt die Ergebnisse in (49) ein, so bekommt man die Rekursionsformel

$$(54) \qquad n_1(r) = 3n_1(r-2) + 3n_1(r-4) + n_1(r-6) \qquad (r \geq 8).$$

Indem man für die Anfangswerte von r die nötigen Anzahlen direkt berechnet, kommt man zu folgendem Ergebnis:

r	1	2	3	4	5	6	7	8	9	10	11 ...
$n_1(r)$	1	0	1	2	4	6	15	24	58	92	223 ...
$n_2(r)$	0	0	1	0	1	2	5	8	19	30	73 ...
$n_3(r) = n_4(r)$	0	1	1	2	5	8	19	30	73	116	281 ...

Unter Heranziehung dieser Übersicht verifiziert man mittels der vorhergehenden Formeln:

$$(55) \quad n_j(r) = 3n_j(r-2) + 3n_j(r-4) + n_j(r-6) \qquad (j = 1, 2, 3, 4;\ r \geq 9).$$

Die nachfolgende Tabelle zeigt, welche Klassen für $r \leq 6$ (im Falle 2 für $r \leq 8$) auftreten. Leere Klassen sind durch Striche bezeichnet. Unbefriedigend

Fall 1: $c = 0$

r	b_1	b_2	b_3	b_4	b_5
1	leere Folge				
2	—				
3	b_1	$3b_1$			
4	b_1	2	$3b_1+1$		
	b_1	1	$3b_1+2$		
5	b_1	$3b_3$	b_3	$3b_1$	
	b_1	1	1	$3b_1+1$	
	b_1	2	3	$3b_1+1$	
	b_1	1	3	$3b_1+2$	
6	b_1	$3b_4+1$	2	b_4	$3b_1$
	b_1	$3b_4+2$	1	b_4	$3b_1$
	b_1	2	2	4	$3b_1+1$
	b_1	2	3	3	$3b_1+1$
	b_1	1	3	3	$3b_1+2$
	b_1	1	4	2	$3b_1+2$

Fall 2: $c = 1$ minimal

r	b_1	b_2	b_3	b_4	b_5	b_6	b_7
1	—						
2	—						
3	1	1					
4	—						
5	1	2	1	1			
6	1	2	3	1	1		
	1	3	2	1	1		
7	1	1	b_3	$3b_3+1$	1	1	
	1	2	3	3	1	1	
	1	2	1	2	1	1	
	1	3	3	2	1	1	
	1	$3b_3+1$	b_3	1	1	1	
8	1	1	b_3	2	$3b_3+2$	1	1
	1	1	b_3	1	$3b_3+3$	1	1
	1	2	3	3	3	1	1
	1	2	4	2	3	1	1
	1	3	2	4	2	1	1
	1	3	3	3	2	1	1
	1	$3b_4+3$	1	b_4	1	1	1
	1	$3b_4+2$	2	b_4	1	1	1

Fall 3: $c = 1$ maximal

r	b_1	b_2	b_3	b_4	b_5
1	—				
2	2				
3	3	2			
4	3	3	2		
	4	2	2		
5	2	b_2	$3b_2+1$	2	
	3	1	2	2	
	3	3	3	2	
	4	3	2	2	
	$3b_2+2$	b_2	1	2	
6	2	b_2	2	$3b_2+2$	2
	2	b_2	1	$3b_2+3$	2
	3	3	3	3	2
	3	4	2	3	2
	4	2	4	2	2
	4	3	3	2	2
	$3b_3+3$	2	b_3	1	2
	$3b_3+4$	1	b_3	1	2

Fall 4: $c = 2$ maximal

r	b_1	b_2	b_3	b_4	b_5
1	—				
2	1				
3	3	1			
4	3	3	1		
	2	4	1		
5	$3b_2+1$	b_2	2	1	
	2	1	3	1	
	3	3	3	1	
	2	3	4	1	
	1	b_2	$3b_2+2$	1	
6	$3b_3+2$	2	b_3	2	1
	$3b_3+3$	1	b_3	2	1
	3	3	3	3	1
	3	2	4	3	1
	2	4	2	4	1
	2	3	3	4	1
	1	b_2	2	$3b_2+3$	1
	1	b_2	1	$3b_2+4$	1

ist, daß hier im Gegensatz zu $Q=2$ auch Folgen auftreten, die im Sinne der seinerzeitigen Definition (I 688) unbrauchbar sind. Für die übrigen berechnen sich die Zahlen b_0, D, b_r nach den Formeln des § 1, die hier nochmals angeführt seien:

$$2b_0 - c = (-1)^{r-1} Q\, B_{r-2} B_{r-3,1} + m\, B_{r-1},$$
$$D - b_0^2 = (-1)^{r-1} Q\, B_{r-3,1}^2 + m\, B_{r-2,1},$$
$$Q\, b_r = 2(b_0 - c).$$

Dabei ist für m jede ganze Zahl zulässig, die noch gewissen Bedingungen (I 687) genügt. Hat man für eine bestimmte Q, c-Folge die Größen b_0, D, b_r schon mittels m berechnet, so kann man die zu irgendeinem anderen Parameter $m' = m + x$ gehörigen Werte b_0', D', b_r' noch bequemer finden als mit Hilfe der vorstehenden Formeln; aus diesen folgt nämlich

$$\text{(56)} \qquad \begin{cases} b_0' - b_0 = \frac{1}{2} x\, B_{r-1}, \\ b_0'^2 - b_0^2 = x\, b_0\, B_{r-1} + \frac{1}{4} x^2 B_{r-1}^2, \end{cases}$$

$$\text{(57)} \qquad \begin{cases} D' - D = b_0'^2 - b_0^2 + x\, B_{r-2,1} \\ \qquad\quad = x(b_0 B_{r-1} + B_{r-2,1}) + \frac{1}{4} x^2 B_{r-1}^2 \\ \qquad\quad = x\, A_{r-1} + \frac{1}{4} x^2 B_{r-1}^2, \end{cases}$$

wobei A_{r-1} mittels b_0 zu berechnen ist[1]);

$$\text{(58)} \qquad b_r' - b_r = \frac{x}{Q} B_{r-1}.$$

Dabei ist wichtig, daß für x nur solche Zahlen in Frage kommen, für welche die Beziehungen

$$\text{(59)} \qquad Q \mid x,\ 2 \mid x\, B_{r-1},\ x\, B_{r-1} > 2(c - b_0)$$

gleichzeitig gelten; das folgt aus den erwähnten Bedingungen, denen m und m' genügen müssen.

Durch den Satz 2 (I 692) wird jeder Klasse des Falles 2 eine ebensolche (unter Umständen die nämliche) zugeordnet. Die Tabelle ist so angelegt, daß in diesem Sinn bei festem r jeweils die erste Zeile der letzten, die zweite der vorletzten usw. entspricht. Der nämliche Satz paart jede Klasse des Falles 3 mit einer solchen des Falles 4; diese Paare stehen jeweils in der nämlichen Tabellenzeile.

[1]) Vgl. Schmidt, Hermann: Zur Approximation und Kettenbruchentwicklung quadratischer Zahlen. Math. Z. **52**, 168—192 (1950); hier S. 179.

(13b) Alzing, Post Bergen 2, Dr. Liegl-Str. 10

(Eingegangen am 6. Juni 1956)

TIETZE, H.
Math. Zeitschr. Bd. 67, S. 188—202 (1957)

Über Schachturnier-Tabellen

Herrn GEORG FABER zum 80. Geburtstag am 5. April 1957 gewidmet

Von

HEINRICH TIETZE

§ 1. Einleitung

Lieber Faber! Was ich Ihnen mit meinen herzlichsten Wünschen zu Ihrem 80. Geburtstag vorlege, sind nur Bruchstücke, zum Teil wohl noch stammend aus der Zeit unserer Schachpartien in einem kleinen Freundeskreis. Es handelt sich um ein bescheidenes kombinatorisches Problem, dessen hier noch nicht vollständig gegebene Lösung (vgl. Nr. 3) vielleicht einer geschickteren Hand unschwer gelingt. Um auch jemandem, der mit dem Wesen von Schachturnieren weniger vertraut ist, verständlich zu sein, wird zu Beginn der folgenden Entwicklungen einiges auseinandergesetzt, was Ihnen seit langem geläufig ist.

1. Beim Schach sind es *zwei* Gegner, die miteinander eine Partie austragen, abweichend zum Beispiel von vielen Kartenspielen, bei denen drei oder vier Spieler an einem Spiel beteiligt sind (über analoge Fragen für solche Spiele vgl. Nr. 5). Dabei ist die Situation der beiden Schachgegner nicht die gleiche und man unterscheidet den, der die *weißen* und den, der die *schwarzen* Steine führt. Wenn nun ein Turnier veranstaltet wird, bei dem p die Anzahl der daran teilnehmenden Spieler ist, so wird es in der Regel so eingerichtet, daß im Laufe des Turniers jeder Spieler mit jedem anderen eine Partie[1]), jeder also insgesamt $p-1$ Partien spielt, wobei außerdem getrachtet wird, daß jeder Spieler innerhalb der Gesamtheit seiner Partien ungefähr gleich viel Partien mit den weißen, wie mit den schwarzen Steinen zu spielen hat.

Zur Erledigung aller dieser $\frac{1}{2}p(p-1)$ Partien müssen die Spieler des öfteren zusammenkommen — eine solche Zusammenkunft wird eine „Runde" genannt — und damit sich das Turnier nicht über eine zu lange Zeit erstreckt, wird man trachten, daß bei einer Runde möglichst viele der p Turnierspieler beschäftigt, also möglichst wenige „spielfrei" sind. Man sucht also, es so

[1]) Gelegentlich sind wohl auch „Doppelrunden-Turniere" vorgekommen, in denen jeder Spieler mit jedem anderen zwei Partien zu spielen hat. Da man es dabei einfach so einrichtet, daß nach einer ersten Turnierhälfte, in der jeder mit jedem eine Partie gespielt hat, nun in einer zweiten Turnierhälfte dieselben Spielerpaarungen, nur mit Vertauschung von Weiß und Schwarz, vorkommen, so ergeben solche Doppelrunden-Turniere keine besonderen, über die gewöhnlichen „Einrunden-Turniere" hinausgehenden Verteilungsprobleme.

einzurichten, daß nach r Runden die Anzahl der Partien, die ein einzelner Spieler gespielt hat, bei möglichst viel Spielern gleich r oder nicht viel kleiner als r ist.

Zu dieser Forderung tritt noch eine, die die Verteilung der Farben Weiß und Schwarz auf die einzelnen Spieler betrifft. Wir sagten schon, daß jeder Spieler innerhalb des gesamten Turniers mit jeder der beiden Farben ungefähr gleich oft spielen soll. Darüber hinaus fordert man, daß bei jedem Spieler auch im Verlauf des Turniers die Zahl der Partien, die er bis dahin mit Weiß, und die Zahl der Partien, die er mit Schwarz gespielt hat, sich dauernd ungefähr die Wage halten; daß also, wenn nach r Runden ein Spieler r (oder vielleicht $r-1$) Partien gespielt hat, er in diesen Partien etwa ebenso oft (also in ungefähr $\frac{1}{2}r$ Partien) Weiß wie Schwarz hatte. Um dies zu erreichen trachtet man, es so einzuteilen, daß beim Übergang von einer Runde zur nächsten möglichst viele Spieler die Farbe der Steine, mit denen sie spielen, wechseln, d.h. daß ein Spieler, der in der einen Runde Schwarz hatte, in der nächsten mit Weiß spielt, und umgekehrt.

2. Allen diesen Forderungen kann in weitgehendem Maße entsprochen werden, wie dies die üblichen „Turnier-Tabellen" zeigen. Um diese Tabellen anzuwenden, findet zu Beginn eines Turniers eine Verlosung statt, bei der die beteiligten Spieler ihre Nummern von 1 bis p bekommen. Es gibt dann die auf die Spielerzahl p bezügliche Turnier-Tabelle für jede Runde an[2]), welche Spieler miteinander zu spielen haben; und wer mit den weißen, wer mit den schwarzen Steinen[3]).

3. In den folgenden Entwicklungen wollen wir nicht von den in der Praxis gebräuchlichen Turnier-Tabellen ausgehen, sondern von den *Forderungen*, die wir bezüglich der Zahl der in den einzelnen Runden mitwirkenden Spieler und bezüglich der Verteilung der Farben auf die p Spieler im Laufe der einzelnen Runden stellen. Diese (in § 2 aufgestellten) Forderungen geben präziser gefaßt das, was in Nr. **1** als ungefähr erstrebenswert geschildert wurde. Natürlich sind wir von vorneherein darauf bedacht, keine Forderungen aufzustellen, die sich nachher als unerfüllbar erweisen würden.

Für das Ergebnis unserer Untersuchungen wird es, wie wir sehen werden, einen Unterschied ausmachen, ob die Anzahl p der Spieler ungerade oder gerade ist: bei ungeradem p ergeben sich im wesentlichen keine anderen unseren Forderungen genügenden Turnier-Verteilungen als die durch die längst gebräuchlichen Turnier-Tabellen gegebenen; bei geradem p wird sich zeigen, daß außer den gebräuchlichen noch andere Verteilungen möglich sind. Und es bleibt hier eine offene Frage (siehe Nr. **22**, **23**), wie die Gesamtheit

[2]) Darüber, daß vormals umständlicher verfahren wurde, vgl. Anm. 6.

[3]) Dagegen wäre es gewöhnlich (nämlich abgesehen von einer ganz kleinen Zahl von Spielern) nicht angängig, zunächst nur für die erste Runde durchs Los die Verteilung der Spieler zu bestimmen, dann durch neues Losen für die zweite Runde usw. Denn es würde sich meistens bald zeigen, daß den in Nr. **1** erörterten Forderungen dann nicht entsprochen werden kann (siehe Beispiele in Nr. **8**).

aller Verteilungen beschaffen ist, die unseren Forderungen bei geradem p genügen.

Das eine aber wird allemal, ob nun p ungerade oder gerade ist, aus unseren Entwicklungen hervorgehen, nämlich daß man *keine schärferen* Forderungen an die Turnier-Verteilungen stellen kann, als diejenigen, die von den üblichen Turnier-Tabellen erfüllt werden.

4. Was solche Turnier-Tabellen betrifft, so ist es selbstverständlich, daß man aus einer brauchbaren Tabelle für p Spieler mit den Nummern 1 bis p wiederum eine brauchbare Tabelle erhält, wenn man die Nummern der Spieler irgend einer Permutation unterwirft. Zwei derart in einander überführbare Tabellen sehen wir als nicht wesentlich verschieden an.

5. Wenn ich nach Zusammenstellung der hier vorgelegten Entwicklungen etwas in der vorhandenen Literatur nachgesehen habe, inwieweit sich da wohl dieselben Ergebnisse vorfinden, so bekam ich den Eindruck, daß es den betreffenden Publikationen vor allem um den Nachweis zu tun ist, daß die mitgeteilten Verfahren geeignet sind, eine brauchbare, allen aufgestellten Forderungen genügende Turnier-Verteilung zu liefern; anders gesagt, daß diese Verfahren zur Lösung des gestellten Problems *hinreichen*. Dazu tritt nun aber die Frage, ob die gewonnene Lösung auch *notwendig* ist, in dem Sinne, daß keine andere allen Forderungen genügende Lösung vorhanden ist. Eine auf diese zweite Frage bezügliche Untersuchung habe ich an den eingesehenen Stellen nicht gefunden[4]). Gerade mit dieser Frage aber befassen sich die folgenden Entwicklungen. Die Antwort, zu der wir gelangen, lautet verschieden, je nachdem ob die Anzahl p der Spieler ungerade ist, wo die übliche Lösung, von einer trivialen Modifikation abgesehen, sich als eindeutig erweist, oder ob p gerade ist (> 2), wo wesentlich verschiedene Lösungen existieren[5]).

Ein geringeres Maß von Forderungen als bei Schachturnieren tritt bei anderen kombinatorischen Aufgaben auf, die in verschiedenen Gestalten gestellt worden sind. Bei diesen Aufgaben (Spaziergänge eines Pensionats, Organisation eines Fechtturniers) entfällt diejenige Unterscheidung zwischen den beiden Elementen eines Paares, die sich beim Schach durch die Zuweisung der weißen bzw. schwarzen Steine für die Spieler einer Partie ergibt.

Ohne Anspruch auf Vollständigkeit seien einige Literaturangaben gebracht. Zunächst für das Problem der Schachturnier-Tabellen: W. AHRENS, Mathematische Unterhaltungen und Spiele, 2. Aufl., Bd. II (1918), Kap. XIV, Abschnitt 1, § 2, S. 79—94. M. KRAITCHIK, La Mathématique des Jeux, ou Récréations Mathématiques, 1930, Chap. 18, Nr. 1—4, S. 540—546. Siehe auch W. AHRENS, Mathematische Spiele, Artikel I G 1 der Encyklopädie der

[4]) Hier möchte allerdings auf seinen Briefwechsel aus 1907 und 1909 mit ROBERT REMAK über die Frage nach verschiedenen Paarungs-Tabellen hingewiesen werden, den W. AHRENS in „Mathematische Unterhaltungen und Spiele", 2. Aufl., Bd. II (1918), S. 92 erwähnt.

[5]) Dies gilt, abweichend von gegenteiligen Behauptungen, schon für $p = 4$ (vgl. Nr. **23**).

math. Wiss., Bd. I, Teil 2 (1900—1904), Anm. 3, S. 1081. Aus der Zeit, wo über Herstellung von Turnier-Tabellen diskutiert[6]) wurde: R. SCHURIG, Deutsche Schachzeitung, 41. Jg., Leipzig 1886, S. 134 u. 169 (in diesem Aufsatz werden die heute vorwiegend benützten Tabellen entwickelt; vgl. Nr. **18**, Anm. 22).

Über allgemeinere Fragen der Aufteilung von Teilnehmern auf Paare vgl. E. LUCAS, Récréations Mathématiques, t. II (1883), S. 176ff.

Über verwandte Fragen speziell über Tripelsysteme (anstelle der Paarungen beim Schach), die von D. P. KIRKMAN und J. STEINER gestellt wurden, siehe Encyklop. d. math. Wiss., I. Bd., I. Teil, Art. I A 2, NETTO, Nr. **10**, S. 33; eine ausführliche Darstellung in W. W. Rouse BALL, Mathematical recreations and problems, 2d ed., 1892, S. 89. Neuere Arbeiten von S. BAYS, TH. SKOLEM, H. HASSE, R. D. CARMICHAEL (1931—1933), dann von S. BAYS und E. DE WECK (Vierersysteme), M. H. MARTIN, H. S. WHITE (1935), ferner von R. C. BOSE (1939) findet man im Zentralblatt, Bde. 1, 2, 3, 5 bzw. 10, 11 bzw. 23 referiert.

Speziell über Verteilungsprobleme für Spiele, bei denen, nicht wie beim Schach zwei, sondern mehr Spieler ein Spiel austragen, vgl. eine Untersuchung über Whist-Turniere von $4n$ Teilnehmern von E. H. MOORE, Amer. Journ. of Math. **18** (1896) S. 290—303; siehe ferner PELTESOHN ROSE, Das Turnierproblem für Spiele zu je dreien, Berlin Diss. 1936 (Referat von BAYS, Zentralblatt, Bd. 13).

§ 2. Die genauen Forderungen an eine Turnier-Tabelle

6. Da die Anzahl der in einer Runde beschäftigten Spieler notwendig gerade ist, so ist klar, daß bei ungeradem $p = 2n - 1$ in jeder Runde eine ungerade Anzahl von Spielern, also mindestens ein Spieler spielfrei bleiben muß, während bei geradem $p = 2n$ in jeder Runde sich eine gerade Anzahl spielfreier Spieler ergibt, soferne nicht überhaupt kein Spieler spielfrei ist.

Wir werden nun an die aufzustellenden Turnier-Tabellen bezüglich spielfreier Spieler die schärfsten denkbaren Forderungen stellen; nämlich:

I a. Bei ungeradem p soll in jeder Runde nur ein einziger Spieler spielfrei sein. Demgemäß sollen in jeder Runde $p - 1$ Spieler beschäftigt, also $\frac{1}{2}(p-1)$ Partien gespielt werden. Um also alle $\frac{1}{2}p(p-1)$ Partien des Turniers zu absolvieren, sind genau p Runden erforderlich; und da jeder der p Spieler $p - 1$ Partien zu absolvieren hat, muß er in allen Runden bis auf eine einzige, in der er spielfrei bleibt, beschäftigt sein; in den einzelnen p Runden ist also der Reihe nach jeder Spieler einmal spielfrei.

[6]) Statt einer einzigen Auslosung wurden beispielsweise bei einem Meisterturnier 1877 zwei Verlosungen, — eine für die Paarungen der Spieler, eine zweite für die Reihenfolge der Runden — vorgenommen, und dies zur Vereinfachung früherer noch umständlicherer Verfahren; siehe E. SCHALLOPP, Der Schachkongreß zu Leipzig 1877, Leipzig 1878, S. 33—34.

Ib. Bei geradem p soll es in keiner Runde einen spielfreien Spieler geben. In jeder Runde sollen also $\frac{1}{2}p$ Partien gespielt werden, und man braucht $p-1$ Runden, um alle $\frac{1}{2}p(p-1)$ Partien des Turniers zu erledigen.

7. Was ferner den Wechsel der weißen und schwarzen Steine bei jedem einzelnen Spieler betrifft, so lassen sich, wie sich zeigen wird, die folgenden Forderungen erfüllen.

IIa. Bei ungeradem p soll jeder Spieler in zwei aufeinanderfolgenden der von ihm gespielten Partien stets die Farbe der Steine, mit denen er spielt, wechseln. Dabei wird, wenn er in einer Runde eine Partie gespielt hat, die nächste von ihm zu spielende Partie im allgemeinen in der nächsten Runde zu spielen sein; und nur dann in der übernächsten Runde, wenn er in der nächsten Runde spielfrei ist.

Aus IIa folgt, daß jeder Spieler in genau $\frac{1}{2}(p-1)$ seiner Partien Weiß, in ebensoviel Partien Schwarz hat.

IIb. Bei geradem p soll es bei jedem Spieler höchstens einmal vorkommen, daß er in zwei aufeinanderfolgenden Partien mit derselben Farbe spielt; im übrigen soll er von jeder Partie zur nächsten stets die Farbe wechseln.

Da die Zahl $p-1$ der Partien, die ein Spieler spielt, ungerade ist, kann er natürlich nicht gleich oft Weiß wie Schwarz bekommen. Aus IIb aber folgt, daß er nur genau einmal mehr mit der einen Farbe als mit der anderen zu spielen hat[7]).

8. In den folgenden Entwicklungen (§§ 3—5) besprechen wir die Konsequenzen, die die aufgestellten Forderungen für die Konstruktion von Turnier-Tabellen nach sich ziehen. Dabei mögen vorher noch einige Beispiele zeigen, daß man, wie schon in Anm. 3 bemerkt, ins Gedränge kommen kann, wenn man ohne einen Gesamtplan für das ganze Turnier, einfach durch schrittweises Verlosen für die einzelnen Runden die Verteilung der Spieler bestimmen wollte.

In einem ersten Beispiel handle es sich um $p=5$ Spieler, wobei wir annehmen, daß zunächst die Reihenfolge verlost wurde, in der die Spieler in den einzelnen Runden spielfrei sind. In dieser Reihenfolge mögen die Spieler selbst mit den Nummern 1, 2, 3, 4, 5 bezeichnet werden. Für die erste Runde habe dann eine Verlosung im einzelnen die Verteilung ergeben:

$$(R_1) \qquad (1),\ 5-2,\ 4-3,$$

also 1 spielfrei, 2 habe 5 zum Gegner bekommen, 3 spielt somit gegen 4; und die Verlosung der Farben habe für 5 und 4 Weiß ergeben. Für die nächste Runde, wo 2 seinen freien Tag hat, möge sich beim Auslosen des Gegners von

[7]) Eine weitere Forderung III, die darauf hinausläuft, Tabellen für ungerade Spieleranzahlen entbehrlich zu machen, wird erst am Schluß (in Nr. 25) kurz erwähnt werden.

1 der Spieler 4 ergeben haben[8]); die andere Partie wird dann zwischen 3 und 5 gespielt, sodaß die zweite Runde — unter Beachtung der Farben, die sich aus Forderung IIa ergeben, — lautet:

(R_2) $\qquad (2),\ 1-4,\ 3-5.$

Für die dritte Runde, in der 3 spielfrei ist, kämen nun als Gegner von 1 nur 2 oder 5 in Frage. Im ersten Fall, wo sich die Partie $2-1$ ergäbe, könnte die andere Partie zwischen 4 und 5 nicht zustandekommen, weil diese Spieler gemäß IIa beide mit Weiß zu spielen hätten. Im zweiten Fall, wo sich die Partie $5-1$ ergäbe, fiele die andere Partie aus, weil 3 und 4 ihre Partie schon in (R_1) absolviert haben. Eine allen Forderungen genügende Weiterführung des Turniers erwiese sich also als unmöglich.

In einem zweiten Beispiel mit $p=6$ Spielern seien auf Grund schrittweiser Verlosungen die drei Runden

$$1-4,\ 6-2,\ 3-5;$$
$$2-1,\ 4-3,\ 5-6;$$
$$1-6,\ 2-3,\ 4-5$$

gespielt worden. Der Forderung IIb gemäß müßte nun der Spieler 2, da er zweimal nacheinander Weiß hatte, in der vierten Runde Schwarz, in der fünften und letzten Runde Weiß bekommen. Ganz dasselbe würde aber für den Spieler 4 gelten. In keiner dieser Runden könnte also die Partie zwischen den Spielern 2 und 4 zustandekommen.

Übrigens können schon für $p=4$ durch planlose Verlosungen Schwierigkeiten entstehen. Falls in der ersten Runde die Spieler 1, 2 und 3, 4 ihre Partien spielten, in der zweiten Runde 1, 3 und 2, 4 und man jedesmal den Spielern die Auslosung der Farben überließ und sich dabei die Partien

$$1-2,\ 4-3;$$
$$1-3,\ 4-2$$

ergaben, dann müßte in der dritten und letzten Runde gemäß Forderung IIb sowohl der Spieler 2 als auch der Spieler 3 die Farbe Weiß bekommen, sodaß die zwischen ihnen noch ausständige Partie nicht zustandekäme; und das Entsprechende gilt für die Partie zwischen 1 und 4, die beide Schwarz bekommen müßten[9]).

[8]) Für diese Auslosung kam 5 als Gegner von 1 überhaupt nicht in Frage, da sonst die andere in dieser Runde zu spielende Partie von 3 und 4 zu spielen gewesen wäre, die ihre Partie schon in (R_1) absolviert haben. Wäre aber 3 als Gegner von 1 ausgelost worden, so hätte man diese Verlosung annullieren müssen, denn die andere Partie wäre dann zwischen den Spielern 4 und 5 auszutragen gewesen, der eine mit den weißen, der andere mit den schwarzen Steinen, während doch jeder, weil er in (R_1) mit Weiß spielte, nunmehr nach Forderung IIa Schwarz bekommen müßte, — dies schon ein Zeichen der Verlegenheiten bei aufs Gerathewohl vorgenommenen Auslosungen!

[9]) Analoge Schwierigkeiten können sich bei etwas größerer Teilnehmerzahl durch planlose Abpaarungen auch dann ergeben, wenn man die auf die Farbenverteilung bezüglichen Forderungen IIa, IIb ganz unberücksichtigt läßt und nur Ia, Ib beachtet, also Verteilungen im Auge hat, wie sie bei Fechtturnieren oder bei der Aufgabe über paarweise geordnete Spaziergänge eines Pensionats (siehe Nr. 5) sich ergeben.

§ 3. Die zulässigen Typen der Farbenverteilung

9. Wir beginnen mit dem Fall einer *ungeraden* Spielerzahl $p=2n-1$ ($n\geq 2$). Betrachten wir der Reihe nach die Runden $r=1, 2, \ldots, p$, so ist gemäß Ia (Nr. 6) jeder Spieler charakterisierbar durch die Nummer $r=i$ der Runde, in der er spielfrei ist. Dann bestehen bezüglich der Farben[10]), mit denen er in den Runden $r\leq i$ spielt, gemäß IIa (Nr. 7) nur zwei Möglichkeiten. Die eine dieser Möglichkeiten liegt vor, wenn er in der Runde $i+1$, also der ersten Runde, nachdem er spielfrei war, die Farbe $+1$ hat, und gemäß IIa dann ebenfalls $+1$ in den Runden $i+3, i+5, \ldots$ ($\leq p$), dagegen die Farbe -1 in den Runden $i+2, i+4, \ldots$ ($\leq p$); ferner hat er gemäß IIa die Farbe -1 in den Runden $i-1, i-3, \ldots$ (≥ 1) und die Farbe $+1$ in den Runden $i-2, i-4, \ldots$ (≥ 1). Der Typus der Farbenverteilung des Spielers sieht dann so aus:

$$(1)\qquad \begin{cases} \text{Runde } \ldots, & i-3, & i-2, & i-1, & i, & i+1, & i+2, & i+3, \ldots \\ \text{Farbe } \ldots & - & + & - & * & + & - & + & \ldots, \end{cases}$$

ein Typus, den wir abgekürzt mit $[i*+]$ bezeichnen wollen[11]). Natürlich entfallen die Runden $r<i$, wenn $i=1$ ist; und es entfallen die Runden $r>i$, wenn $i=p$ ist.

Die zweite der genannten Möglichkeiten liegt vor, wenn der Spieler, der in der Runde $r=i$ spielfrei ist, in der Runde $i+1$ die Farbe -1, dagegen in der Runde $i-1$ die Farbe $+1$ hat. Seine Farbenverteilung ergibt sich aus (1), indem man durchwegs $+$ und $-$ vertauscht. Der so gekennzeichnete Typus werde mit $[i*-]$ bezeichnet.

10. Bei *geradem* $p=2n$ ($n\geq 2$) betrachten wir zunächst die Typen der Farbenverteilung für diejenigen Spieler, die einmal öfter mit Weiß ($+$) als mit Schwarz ($-$) spielen, also in $\frac{1}{2}p=n$ Runden Weiß haben und in $\frac{1}{2}p-1=n-1$ Runden Schwarz. Ein solcher Typus liegt speziell dann vor, wenn der Spieler in der ersten Runde $+$ hat und dann dauernd die Farbe wechselt, so daß seine Farbenverteilung so aussieht:

$$(2)\qquad \begin{cases} \text{Runde} & 1, & 2, & 3, \ldots, & p-2, & p-1 \\ \text{Farbe} & + & - & + \cdots & - & + \,. \end{cases}$$

Ferner gehören hierher die Fälle, wo der Spieler in zwei aufeinanderfolgenden Runden, etwa $r=i$ und $r=i+1$ ($1\leq i\leq p-2$) mit $+$ spielt und im übrigen

[10]) Zur Kennzeichnung der Farbe Weiß schreiben wir zur Abkürzung $+$, für Schwarz analog $-$. Statt $+$ und $-$ schreiben wir auch $+1$ und -1; das wird sich speziell dann als bequem erweisen, wenn wir später gelegentlich Aussagen zu machen haben, wie die, daß ein Spieler mit Weiß oder mit Schwarz spielt, je nachdem ob eine auftretende Zahl m gerade oder ungerade ist; wir können dann einfach sagen, er spielt mit der Farbe $(-1)^m$.

[11]) Dieses Zeichen mag im Hinblick auf (1) so gedeutet werden, daß bei i ein Stern $*$ steht, darauf dann ein $+$ folgt; ausführlicher könnte man dafür $[i*, (i+1)+]$ schreiben.

fortwährend die Farbe wechselt, was das Bild gibt:

$$(3)\quad \begin{cases} \text{Runde } \ldots,\ i-2,\ i-1,\ i,\ i+1,\ i+2,\ i+3,\ i+4,\ \ldots \\ \text{Farbe } \ldots\ +\quad -\quad +\quad +\quad -\quad +\quad -\quad \ldots, \end{cases}$$

wobei natürlich, wenn $i=1$ bzw. $i=p-2$ ist, die Runden $r<i$ bzw. $r>i+1$ wegfallen.

Der Typus der durch (3) gekennzeichneten Farbenverteilung werde zur Abkürzung mit $[i+]$ bezeichnet[12]; ferner mit $[(p-1)+]$ der Typus, der durch (2) dargestellt wird[13].

Aus den soeben besprochenen Typen, wo der Spieler einmal öfter + als − hat, ergeben sich die Typen, wo er einmal öfter − als + hat, durch Vertauschung von + mit −. Diese Typen werden wir abgekürzt mit $[i-]$ bezeichnen $(1 \leq i \leq p-1)$.

§ 4. Turnier-Verteilungen bei ungeradem p

11. In Nr. **9** haben wir mit den Spielertypen $[i*+]$ und $[i*-]$ $(i=1, 2, \ldots, p)$ insgesamt $2p$ mögliche Typen festgestellt. Bei einem Turnier können natürlich niemals zwei Spieler den gleichen Typus aufweisen, da es ja eine Runde geben muß, in der sie gegen einander zu spielen haben, somit verschiedene Farben aufweisen. Die p Spieler werden also p verschiedene Typen repräsentieren: In einem Turnier tritt eine bestimmte *Kombination von p Spielertypen* auf. Unsere erste Aufgabe wird es sein, festzustellen, welche Kombinationen von p Typen in einem und demselben Turnier möglich sind. Nachher (Nr. **13**—**17**) wird — als zweite Aufgabe — für eine solche Kombination von Typen zu ermitteln sein, welche Paarungen der Spieler in den einzelnen Runden mit einer solchen Kombination in Einklang stehen; und ob es zu einer solchen Kombination nur ein einziges System von Paarungen der Spieler gibt, oder ob bei gleicher Typen-Kombination mehrere „Turnier-Verteilungen" möglich sind. Es wird sich zeigen, daß nur *eine*, durch die Kombination der Spielertypen also eindeutig bestimmte Turnier-Verteilung existiert[14].

12. Da gemäß Nr. **6** (Forderung Ia) und **9** in jeder Runde $r=i$ genau ein Spieler spielfrei ist, so ist unter den Spielern für jedes i von den beiden Typen $[i*+]$ und $[i*-]$ genau einer vertreten.

Wir vergleichen nun den Typus $[i*+]$ mit $[(i+1)*-]$, der eine dargestellt durch (1), der andere durch:

$$(4)\quad \begin{cases} \text{Runde } \ldots,\ i-3,\ i-2,\ i-1,\ i,\ i+1,\ i+2,\ i+3,\ \ldots \\ \text{Farbe } \ldots\ -\quad +\quad -\quad +\quad *\quad -\quad +\quad \ldots. \end{cases}$$

---

[12] Im Hinblick auf den Anblick von (3) wäre $[i+, (i+1)+]$ eine ausführlichere Kennzeichnung für diesen Typus; vgl. Anm. 11.

[13] Man mag hierbei an eine zyklische Anordnung der Runden denken, so daß auf die Runde $i=p-1$ die Runde $i=1$ folgt, was man auch so auffassen kann, daß man die Nummern der Runden mod. $p-1$ betrachtet.

[14] Verglichen mit diesen einfachen Verhältnissen bei ungeradem p herrscht bei geradem p eine wesentlich größere Mannigfaltigkeit (vgl. § 5, Nr. **23**, **24**).

Man sieht nun sofort:

Satz 1. Es ist ausgeschlossen, daß in einem Turnier sowohl der Spieler vom Typus $[i*+]$ als auch der vom Typus $[(i+1)*-]$ vorkommt.

Denn in irgend einer Runde müßten die beiden Spieler miteinander eine Partie austragen, der eine mit $+$, der andere mit $-$. Die Runden $r=i$ und $r=i+1$ scheiden dafür aus, weil hier jedesmal einer der beiden Spieler spielfrei ist. In allen anderen Runden, also $r<i$ und $r>i+1$ stimmen aber die beiden Spieler, wie der Vergleich von (1) mit (4) zeigt, in den Farben überein, können einander also nicht in einer Partie begegnen[15]).

Genau so ergibt sich, daß in einem Turnier nicht sowohl der Typus $[i*-]$ als auch der Typus $[(i+1)*+]$ vorkommen kann. Da nun nach dem oben Gesagten von den beiden Typen $[(i+1)*+]$, $[(i+1)*-]$ notwendig einer vorkommen muß, so sehen wir, daß zugleich mit $[i*+]$ nur $[(i+1)*+]$, zugleich mit $[i*-]$ nur $[(i+1)*-]$ vorkommen kann. Wenn man das der Reihe nach von $i=1$ bis $i=p-1$ beachtet, ergibt sich:

Satz 2. Bei einem Turnier mit ungerader Spieleranzahl p sind keine anderen Kombinationen von Spielertypen möglich, als die aus den Typen

$$[1*+],\ [2*+],\ \ldots,\ [p*+] \tag{5}$$

und die aus den Typen

$$[1*-],\ [2*-],\ \ldots,\ [p*-] \tag{6}$$

bestehende.

Wir wollen eine Kombination von Spielertypen *„realisierbar"* nennen, wenn es wenigstens eine den Forderungen von § 2 genügende Verteilung aller $\frac{1}{2}p(p-1)=(2n-1)(n-1)$ Partien zwischen den p Spielern gibt, welche diese Kombination von Typen aufweist. Satz 2 kann dann so ausgesprochen werden:

Satz 3. Realisierbar kann keine von (5) und (6) verschiedene Kombination sein.

Da wir nun tatsächlich Partie-Verteilungen werden auffinden können (Nr. 13—17), die mit (5) bzw. (6) in Einklang stehen, so wird damit die erste der in Nr. 11 gestellten Aufgaben gelöst sein.

13. Wir betrachten den ersten Fall[16]), also den durch die Typen (5) dargestellten. Die folgenden Entwicklungen werden nun nicht nur die Realisierbarkeit von (5), also die Existenz einer (5) aufweisenden Turnier-Verteilung zeigen, sondern darüber hinaus, daß für alle Runden $r=1, 2, \ldots, p$ die Turnier-

[15]) Dasselbe gilt, wie man sich leicht überzeugt, für $[p*+]$ und $[1*-]$. Man könnte diese Aussage dem Satz 1 einordnen, wenn man festsetzt, daß für $i=p$ unter $i+1=p+1$ nichts anderes als 1 verstanden werden soll. Die hierin auftretende Erscheinung, daß die vorkommenden Zahlen modulo p zu betrachten sind, werden wir später noch öfter antreffen.

[16]) Auf diesen wird sich der zweite Fall mit den Typen (6) ohne weiteres zurückführen lassen (Nr. 17).

Verteilung — also wie die Spieler gepaart sind und mit welcher Farbenverteilung — eindeutig durch unsere Forderungen festgelegt ist. Dabei werden uns gewisse Vorstellungen dienlich sein, die wir jetzt einführen wollen.

Zunächst betrachten wir eine Kreislinie L, die in p gleiche Teile geteilt sei („Elementarbogen" von der Bogenlänge $2\pi:p$), so daß die Teilungspunkte die Ecken eines regelmäßigen p-Ecks bilden. Nach Wahl eines positiven Umlaufssinnes (etwa den im Sinn des Uhrzeigers) denken wir uns den p Ecken der Reihe nach die Nummern $1, 2, \ldots, p$ gegeben, wobei nun 1 als der auf p nächstfolgende Punkt erscheint[17]). Irgend zwei Eckpunkte, etwa a und b, zerlegen nun den Kreis in zwei ungleiche Teilbogen, von denen der eine, den wir den „inneren" Bogen nennen, eine gerade Anzahl $2h$ von Elementarbogen umfaßt ($1 \leq h \leq n-1$), der andere — der „äußere" Bogen — eine ungerade Anzahl $p-2h$ (wo $1 \leq p-2h \leq p-2$). Denken wir den ersteren, also den inneren Bogen, in positivem Sinn durchlaufen, und sei i die Nummer des Anfangspunktes (wo entweder $i=a$ oder $i=b$ ist), so kann der Endpunkt (also der andere der beiden Punkte a, b) mit $i+2h$ bezeichnet werden, soferne man diese Bezeichnung mod. p nimmt (also $i+2h$, falls $>p$, durch $i+2h-p$ ersetzt). Von Interesse ist dann noch, wenn wir die Sehne zwischen den Punkten a, b und den auf ihr senkrechten Kreisdurchmesser betrachten, derjenige Eckpunkt unseres Polygons, der auf diesem Durchmesser liegt. Es ist das der Punkt, der den inneren Bogen halbiert, und es kommt ihm die Nummer $i+h$ zu[18]). Das Ergebnis der folgenden Überlegungen wird dann sein:

Satz 4. Wenn für zwei Spieler $[a*+]$, $[b*+]$ in der eben geschilderten Weise die Ecken mit den Nummern a, b auf L einen inneren Bogen mit $2h$ Elementarbogen bilden, sodaß diese Nummern durch i und $i+2h$ bezeichnet werden können[19]), dann findet die Partie zwischen diesen Spielern notwendig in der Runde $r=i+h$ statt[20]), wobei der Spieler $[i*+]$ mit der Farbe $(-1)^{h-1}$, der Spieler $[(i+2h)*+]$ mit $(-1)^h$ spielt.

Dieser Satz besagt beispielsweise für $p=9$ Spieler, daß die Spieler $[5*+]$, $[9*+]$, wo $i=5$, $i+2h=9$, $h=2$, $i+h=7$ ist, in der Runde $r=7$ miteinander spielen, der Spieler $[5*+]$ mit der Farbe $(-1)^{h-1}=-1$, der Spieler $[9*+]$ mit der Farbe $(-1)^h=+1$; wenn wir dagegen (wieder für $p=9$) unseren Satz anwenden auf die Spieler $[6*+]$ und $[9*+]$, wo jetzt der innere Bogen bei $i=9$ anfängt und mit $h=3$ zum Punkt $i+2h=15$ führt, der nichts anderes ist als $15-p=6$, und wo jetzt $r=i+h=12$ (modulo $p=9$) durch 3 zu ersetzen ist, dann lehrt Satz 4, daß die beiden Spieler in der Runde $r=3$ miteinander spielen, dabei der Spieler $[i*+]=[9*+]$ mit der Farbe

---

[17]) Anders gesagt betrachten wir die Nummern der Punkte modulo p, so daß 1 zugleich als Vertreter der Nummer $p+1$ auftritt.

[18]) Es ist das wieder — im Sinn einer Auffassung mod. p — so zu verstehen, daß man $i+h$, falls $>p$ durch $i+h-p$ ersetzt.

[19]) Wie oben besprochen ist $i+2h$, falls $>p$, durch $i+2h-p$ zu ersetzen.

[20]) Bezüglich $i+h$ gilt wieder das in Anm. 18 Gesagte.

$(-1)^{h-1}=+1$, der Spieler $[6*+]=[(i+2h)*+]=[15*+]$ mit der Farbe $(-1)^h=-1$.

14. Ehe wir den Beweis von Satz 4 (durch vollständige Induktion bezüglich h) erbringen, betrachten wir zunächst für zwei Spieler vom Typus $[a*+]$ und $[b*+]$ sowohl den durch die Punkte a, b auf dem Kreis L bestimmten „inneren Bogen" mit $2h$ Elementarbogen, als auch den ihn zum vollen Kreis L ergänzenden „äußeren Bogen", wobei i und $i+2h$ (natürlich mod. p verstanden) keine anderen Punkte bedeuten als (in geeigneter Reihenfolge) a und b. Wir sehen uns dann für die einzelnen Runden die Farben F_i und F_{i+2h} an, die dem Spieler $[i*+]$ und dem Spieler $[(i+2h)*+]$ in jeder dieser Runden zugewiesen sind. Dies mag durch

$$(7)\quad \begin{cases} \text{Runde } \ldots,\ i-2,\ i-1,\ i,\ i+1,\ \ldots,\ i+h,\ \ldots,\ i+2h-1,\ i+2h,\ i+2h+1,\ \ldots \\ F_i \quad \ldots\ +\ -\ *\ +\ \ldots\ (-1)^{h-1}\ \ldots\ +\ -\ +\ \ldots \\ F_{i+2h} \quad \ldots\ +\ -\ +\ -\ \ldots\ (-1)^{h}\ \ldots\ -\ *\ +\ \ldots \end{cases}$$

dargestellt werden. Dabei sollte man eigentlich an eine zyklische Anordnung der Runden denken, deren Nummern also den in Nr. **13** erwähnten Eckpunkten eines dem Kreis L eingeschriebenen regulären p-Ecks zugeordnet seien. Dann sind durch eine Zahl r mit $i<r<i+2h$ die Eckpunkte gekennzeichnet, die auf dem „inneren Bogen" liegen, während etwa vorhandene[21]) Eckpunkte auf dem „äußeren Bogen" teils durch $r>i+2h$ teils durch $r<i$ charakterisiert sind. Da nun, wie man aus (7) ersieht, für die Runden, deren Nummern zu Eckpunkten auf dem äußeren Bogen gehören, die Farben der beiden Spieler übereinstimmen, so scheiden diese Runden für eine Partie zwischen diesen Spielern aus, desgleichen natürlich die Runden $r=i$ und $r=i+2h$, in denen je einer der beiden Spieler spielfrei ist. Somit gilt

Satz 5. Für eine Partie zwischen den Spielern $[i*+]$ und $[(i+2h)*+]$ kommt nur eine Runde in Betracht, deren Nummer r dem „inneren Bogen" angehört, der auf L durch die Ecken i und $i+2h$ bestimmt ist.

15. Für den speziellen Fall $h=1$, der durch

$$(8)\quad \begin{cases} \text{Runde } \ldots,\ i-2,\ i-1,\ i,\ i+1,\ i+2,\ i+3,\ \ldots \\ F_i \quad \ldots\ +\ -\ *\ +\ -\ +\ \ldots \\ F_{i+2} \quad \ldots\ +\ -\ +\ -\ *\ +\ \ldots \end{cases}$$

dargestellt wird, wo auf dem inneren Bogen nur der eine Eckpunkt mit der Nummer $i+1$ liegt, ergibt sich aus Satz 5 sofort

Satz 6. Die Partie zwischen den Spielern $[i*+]$ und $[(i+2)*+]$ findet notwendig in der Runde $r=i+1$ statt, wobei der Spieler $[i*+]$ die Farbe $+1$, der Spieler $[(i+2)*+]$ die Farbe -1 hat. Im Sinne einer zyklischen

[21]) Sie fallen fort, falls $h=n-1$, $i+2h=i-1+p=i-1$ ist.

Anordnung ist bei $i=p-1$ dabei unter $[(i+2)*+]$ der Spieler $[1*+]$ zu verstehen; ferner bei $i=p$ unter $[(i+2)*+]$ der Spieler $[2*+]$, wobei außerdem $r=i+1$ durch $r=1$ zu ersetzen ist.

16. Mit Satz 6 ist die Gültigkeit von Satz 4 für $h=1$ erwiesen. Um nun Satz 4 durch vollständige Induktion nachzuweisen, nehmen wir an, es sei $2\leq k\leq n-1$ und für alle Werte h, mit $1\leq h\leq k-1$, sei die Gültigkeit von Satz 4 erkannt. Diese Annahme besagt dann, daß der Spieler $[i*+]$ der Reihe nach in den Runden $r=i+1, \ldots, r=i+h, \ldots, r=i+k-1$ mit den Spielern $[(i+2)*+], \ldots, [(i+2h)*+], \ldots, [(i+2k-2)*+]$ zu spielen hat; desgleichen besagt unsere Annahme, daß die Spieler $[(i+2k-2)*+], \ldots, [(i+2k-2h)*+], \ldots, [(i+2)*+]$ der Reihe nach in den Runden $r=i+2k-1, \ldots, r=i+2k-h, \ldots, r=i+k+1$ mit dem Spieler $[(i+2k)*+]$ zu spielen haben. Von den Runden $i<r<i+2k$, die gemäß Satz 5 allein für eine Partie zwischen den Spielern $[i*+]$ und $[(i+2k)*+]$ in Frage kommen, scheiden also die Runden $i+1\leq r\leq i+k-1$ aus, weil in ihnen $[i*+]$ mit anderen Spielern als $[(i+2k)*+]$ zu spielen hat; desgleichen die Runden $i+k+1\leq r\leq i+2k-1$ wegen der Beanspruchung von $[(i+2k)*+]$ durch andere Gegner als $[i*+]$. Es bleibt also für die Partie von $[i*+]$ mit $[(i+2k)*+]$ nur die Runde $r=i+k$ übrig, natürlich mit den Farben $(-1)^{k-1}$ bzw. $(-1)^k$, die den beiden Spielern in dieser Runde zugewiesen sind.

Damit ist die Gültigkeit von Satz 4 auch für $h=k$ erkannt und dieser Satz durch vollständige Induktion bewiesen.

17. Mit Satz 4 ist sowohl die Realisierbarkeit der Typen-Kombination (5) als auch die Eindeutigkeit der zugehörigen Turnier-Verteilung festgestellt. Werden in allen Überlegungen durchwegs die Farben + und − miteinander vertauscht, so erhält man das analoge Resultat für die Typen-Kombination (6), sodaß die in Nr. **11** gestellten Aufgaben gelöst sind.

18. Aus den voranstehenden Entwicklungen, speziell aus Satz 4, kann man für irgend eine ungerade Spieleranzahl p (≥ 3) eine Turnier-Tabelle gewinnen, die z. B. für $p=5$, wenn wir den Spieler vom Typus $[i*+]$ kurz mit der Nummer i bezeichnen, so aussähe:

(1), 3 − 4, 5 − 2

(2), 4 − 5, 1 − 3

(3), 5 − 1, 2 − 4

(4), 1 − 2, 3 − 5

(5), 2 − 3, 4 − 1;

also: in der ersten Runde ist 1 spielfrei und es werden die Partien 3 gegen 4, 5 gegen 2 gespielt (dabei 3 und 5 mit Weiß) usw.

In der Praxis sind allerdings meist etwas andere von Richard SCHURIG (l. c. Nr. 5) aufgestellte Tabellen in Gebrauch, die sich aus den hier gewonnenen durch eine Umnumerierung der Spieler ergeben, also (vgl. Nr. 4) keine prinzipiell verschiedenen Tabellen darstellen[22]).

§ 5. Bemerkungen über Turnier-Verteilungen bei geradem p

19. Bei geradem p treten an die Stelle von Ia, IIa die Forderungen Ib, IIb (§ 2). In der Frage nach allen mit diesen Forderungen verträglichen Turnier-Verteilungen bin ich aber nicht zu einem so vollständigen Ergebnis wie für ungerades p gekommen: Man findet jedenfalls eine größere Mannigfaltigkeit zulässiger Turnier-Verteilungen, deren erschöpfende Kennzeichnung ich aber nicht geben kann, sodaß ich mich hier — schon mit Rücksicht auf den sonst zu großen Umfang dieses Beitrags — unter Verzicht auf Beweise auf einige Bemerkungen beschränken will.

20. In Nr. **10** haben wir von Spielern gesprochen, die einmal öfter mit $+$ als mit $-$ spielen, also unter allen $p-1=2n-1$ Runden insgesamt n-mal mit $+$ und $(n-1)$-mal mit $-$; es sind das Spieler, denen ein Typus $[i+]$ für die Farbenverteilung zukommt, wobei i eine der Zahlen $1, 2, \ldots, p-1$ ist[23]). Ein erstes Resultat, das man beweisen kann, ist

Satz 7. Wenn in einem Turnier ein bestimmter Typus $[i+]$ bei einem Spieler vorkommt, dann kommt auch der Typus $[i-]$ vor.

Die Typen treten also in Paaren einander „entgegengesetzter" Typen auf. Ein solches Paar werde kurz mit $[i\pm]$ bezeichnet.

21. Eine Kombination von Typenpaaren

$$(9)\qquad \begin{cases} [i_1\pm],\ [i_2\pm],\ \ldots,\ [i_n\pm] \\ (1\leq i_1<i_2<\cdots<i_n\leq 2n-1) \end{cases}$$

kann man sich veranschaulicht denken dadurch, daß man auf dem Kreis L (vgl. Nr. **13**), auf dem die Punkte $1, 2, \ldots, p-1$ die Ecken eines regelmäßigen $(p-1)$-Ecks bilden, die n Punkte $i_1, i_2, \ldots, i_n$ markiert. Bei diesen Figuren

$$(10)\qquad (i_1, i_2, \ldots, i_n)$$

wollen wir zwei Arten durch besondere Benennungen unterscheiden: wir wollen eine solche Figur „*dicht*" nennen, wenn sie (gegebenenfalls im Sinne

[22]) Ein Prinzip zur bequemen Herstellung der SCHURIGschen Tabellen (das sich gegenüber dem ursprünglichen SCHURIGschen Verfahren durch größere Einfachheit auszeichnet) hat mir (zusammen mit der Ergänzung für gerade Spieleranzahl) vor vielen Jahren gelegentlich eines Zusammentreffens in Göttingen der Schachmeister PRCZEPIORKA mitgeteilt (auf den auch M. KRAITCHIK, l. c. Nr. **5**, S. 544, verweist). Wegen einer geometrischen Deutung für diesen Mechanismus vgl. Süddeutsche Schachzeitung, 1955, Heft 12, S. 276 und Archimedes, Schriftleit. F. Denk, Erlangen, 8. Jahrg., Heft 3/4, 1956, S. 41.

[23]) Natürlich können niemals zwei Spieler den gleichen Typus aufweisen, da ja dann in keiner Runde eine Partie zwischen ihnen möglich wäre, wo doch in einer Partie stets der eine mit $+$, der andere mit $-$ spielt.

zyklischer Anordnung) drei unmittelbar aufeinanderfolgende Eckpunkte enthält; andernfalls „*undicht*". Bei den letzteren kommen also nur vor: teils *isolierte* Eckpunkte i, wo weder $(i-1)$ noch $(i+1)$ zur Figur gehört; teils *Paare* benachbarter Eckpunkte[24]), niemals aber *Tripel* von unmittelbar aufeinanderfolgenden.

Indem wir die eben bei den Figuren (10) eingeführte Unterscheidung auf die Kombinationen von Typenpaaren (9) übertragen, nennen wir auch eine solche Kombination dicht bzw. undicht, wenn es die darstellende Figur ist.

Ferner wollen wir, analog wie in Nr. **12** eine Kombination (9) von Typenpaaren dann „realisierbar" nennen, wenn es wenigstens eine den Forderungen von § 2 genügende Verteilung aller $\frac{1}{2}p(p-1)=n(2n-1)$ Partien zwischen den $2n$ Spielern gibt, welche die Typen-Kombination (9) aufweist.

22. Man kann dann beweisen

Satz 8. Nur undichte Kombinationen von Typenpaaren können realisierbar sein.

Hier bleibt (für beliebiges p) die Frage offen, ob eine Kombination von Typenpaaren außer der Einschränkung, daß sie nicht dicht sein darf, noch weiteren Einschränkungen unterliegt, oder ob jede undichte Kombination auch realisierbar ist.

23. Eine spezielle realisierbare Kombination von n Typenpaaren ist

$$[1\pm],\ [3\pm],\ \ldots,\ [(p-1)\pm]; \tag{11}$$

in dieser sind, gemäß der in Nr. **21** besprochenen zyklischen Anordnung, $[(p-1)\pm]$ und $[1\pm]$ als benachbarte Typenpaare anzusehen. Es ist die Kombination (11) dadurch von Interesse, daß sie den üblichen Turnier-Tabellen entspricht. Aber Beispiele zeigen (und es gibt sie schon für $p=4$, $n=2$), daß bei geradem $p=2n$ durch eine realisierbare Kombination von n Typenpaaren keineswegs allgemein die Turnier-Verteilung eindeutig festgelegt ist (dies ein wesentlicher Unterschied gegenüber dem Fall eines ungeraden p).

24. Noch in anderer Hinsicht zeigt sich bei geradem p — verglichen mit ungeradem p — ein größerer Reichtum an Möglichkeiten: außer der Kombination (11) von Typenpaaren (und solchen, die in trivialer Weise daraus herleitbar sind) erweisen sich noch wesentlich andere Kombinationen als realisierbar, wozu beispielsweise für $p=8$, $n=4$ die Kombination

$$[1\pm],\ [2\pm],\ [4\pm],\ [5\pm] \tag{12}$$

angeführt sei[25]).

[24]) Die n zur Figur gehörenden Eckpunkte des $(2n-1)$-Ecks können natürlich nicht durchwegs isoliert sein. Die Figur muß also mindestens ein Paar benachbarter Eckpunkte enthalten. Es ist das z. B. bei der Figur $(1, 3, \ldots, 2n-1)$ das (im Sinn zyklischer Anordnung) benachbarte Paar $2n-1, 1$.

[25]) Auf Angabe einer Turnier-Verteilung, die diese Typen-Kombination realisiert, sei aus Gründen der Raumersparnis verzichtet.

25. In der großen Mannigfaltigkeit von Turnier-Tabellen, die sich im Rahmen der Forderungen aus § 2 bei gerader Spielerzahl als möglich erweisen, ergibt sich eine merkliche Einschränkung, wenn man noch die folgende Forderung III hinzunimmt, die vom Standpunkt der Praxis von Bedeutung ist, weil sie es gestattet, sich auf Tabellen mit geradem p zu beschränken:

III. In einer Tabelle für $p = 2n$ Spieler soll es wenigstens einen Spieler geben, nach dessen Weglassung eine unseren Forderungen Ia, IIa genügende Tabelle für $2n-1$ Spieler übrig bleibt.

Gemeint ist damit, daß in der so entstehenden Tabelle in jeder Runde derjenige Spieler spielfrei ist, der mit dem weggelassenen Spieler eine Partie zu spielen gehabt hätte. Dabei möge (gegebenenfalls nach einer Umnumerierung der $2n$ Spieler) dem wegzulassenden Spieler die Nummer $2n$ zugewiesen sein, sodaß die übrigbleibenden die Nummern 1 bis $2n-1$ behalten.

Eine nähere Untersuchung zeigt, daß durch Forderung III alle diejenigen Kombinationen von Typenpaaren (9) ausgeschieden werden, bei denen die zugehörige Figur (10) auf dem Kreis L mehr als ein Paar benachbarter Punkte aufweist. Beispielsweise werden durch Forderung III Turnier-Tabellen ausgeschieden, die zur Kombination (12) gehören. Es bleiben dann, von unwesentlichen Abänderungen abgesehen, nur solche Kombinationen wie (11) übrig.

Diese Hinweise mögen hier genügen. Mag sein, daß sich ein andermal Gelegenheit ergibt, darauf zurückzukommen.

München 23, Trautenwolfstr. 7

(Eingegangen am 10. April 1956)

BEHREND, F. A.
Math. Zeitschr. Bd. 67, S. 203—210 (1957)

Uniformizability and compactifiability of topological spaces

By

F. A. BEHREND

In the following we give some necessary and sufficient conditions for the uniformizability of a topological space. The criteria (§ 2) and the proofs (§ 3) remain within the topological space S; no appeal is made to mappings of S into other spaces; in particular, no use is made of the real numbers. As a topological space is compactifiable if and only if it is a uniformizable T_0-space, each criterion obtained, together with the condition that S be a T_0-space, gives a criterion for compactifiability. The compactification can be carried out either by applying the classical Cauchy completion procedure[1]) to S provided with a totally bounded uniform structure as obtained in § 3.3 or 3.6[2]), or by P. Samuel's construction[3]), or, more directly, by an adaptation of the process used by A. Gerolini[4]) which is sketched in § 4. The compactifiability of regular spaces with a normal base as studied by K. Fan and N. Gottesman[5]) is a simple consequence of criterion (II) (§ 4, note; of the three properties of a normal base, only one is essential).

§ 1. Notations

S is a topological space; x, y, z denote points, A, B, X, Y, Z subsets of S, $\bar{A}$ the closure of A, $\mathring{A}$ its interior, $\complement A$ its complement. Two sets A, B are compatible if $A \cap B \neq 0$; a set A and a filter φ are compatible if A is compatible with every $X \in \varphi$; two filters φ, ψ are compatible if every $X \in \varphi$ is compatible with every $Y \in \psi$. An open filter is a filter with a basis of open sets.

By a *string* σ will be understood a family of subsets of S with the following properties:

(i) if $Z \in \sigma$ then Z is either open or closed,

[1]) See, e.g., BOURBAKI, N.: Éléments de Mathématique, Première Partie, Livre III, Chap. II, § 3, no. 5. Paris 1951.

[2]) See also BEHREND, F. A.: Note on the compactification of separated uniform spaces. Akad. Wetensch. Proc. Ser. A59 = Indagationes Math. **18**, 269—270 (1956).

[3]) Ultrafilters and compactification of uniform spaces. Trans. Amer. Math. Soc. **64**, 100—132 (1948).

[4]) Compactification des espaces séparés. C. R. Acad. Sci. Paris **232**, 1056—1058 (1951).

[5]) On compactifications of FREUDENTHAL and WALLMAN. Nederl. Akad. Wetensch. Proc. Ser. A **55** = Indagationes Math. **14**, 504—510 (1952).

(ii) σ is fully ordered with respect to the inclusion relation $\subset$; it has a least element A and a greatest element B; A is closed and B is open;

(iii) if $A_1 \in \sigma$ is closed, $B_1 \in \sigma$ open, and $A_1 \subset B_1$, then there exists a closed $A_2 \in \sigma$ and an open $B_2 \in \sigma$ such that $A_1 \subset B_2 \subset A_2 \subset B_1$.

We note the following consequences of this definition:

(iv) if σ is a string with least element A and greatest element B, then the family σ' of all complements $\mathbf{C}Z$ of the elements Z of σ is a string with least element $\mathbf{C}B$ and greatest element $\mathbf{C}A$;

(v) if σ is a string, $A_1 \in \sigma$ closed, $B_1 \in \sigma$ open, and $A_1 \subset B_1$, then the family σ' of all $Z \in \sigma$ with $A_1 \subset Z \subset B_1$ is a string with least element A_1 and greatest element B_1.

D, F, G, H, U, V, W denote subsets of $S \times S$; $F(x)$ is the set of all y with $(x, y) \in F$; $F(X) = \bigcup_{x \in X} F(x)$. We note the following

Lemma. For any $X \subset S$ and any symmetrical $F \subset S \times S$, $F(\mathbf{C}F(X)) \subset \mathbf{C}X$ [6].

§ 2. Criteria for uniformizability

It will be shown that the following four conditions (I)—(IV) are equivalent:

(I) *To any open neighbourhood $N(x)$ of any point $x \in S$ there exists a string σ of neighbourhoods of x with greatest element $N(x)$.*

Note. This condition replaces Bourbaki's axiom (O_{IV}) [7] which characterizes uniformizability in terms of continuous real functions on S.

(II) *There exists a relation $<$ between subsets of S with the following properties:*

(II.1) if $A < B$ then $A \subset \mathring{B}$,

(II.2) if $A < B$ then there exist A_1, B_1 such that $A < B_1 \subset A_1 < B$,

(II.3) to any open neighbourhood $N(x)$ of any $x \in S$ there exists a neighbourhood $M(x)$ such that $M(x) \leqslant N(x)$,

(II.4) if $A < B$ then $\mathbf{C}B < \mathbf{C}A$.

Note 1. The following is a consequence of (II):

(II.5) if $A < B$ then $\bar{A} \subset \mathring{B}$.

Proof. By (II.2), $A < B_1 \subset A_1 < B$; by (II.1) and (II.4), $A_1 \subset \mathring{B}$, and $\mathbf{C}B_1 < \mathbf{C}A$, $\mathbf{C}B_1 \subset \mathring{\mathbf{C}}\bar{A} = \mathbf{C}\bar{A}$, $\bar{A} \subset B_1$.

Note 2. (II) is equivalent to

(II*) *There exists a relation $<$ between subsets of S such that*

(II*.1) if $A < B$, then A is closed, B is open and $A \subset B$,

(II*.2) = (II.2),

(II*.3) = (II.3).

Proof. 1. Assume a relation $<$ satisfying (II) and define $A \ll B$ if and only if X, Y exist such that $X < Y$ and $A = \bar{X}$, $B = \mathring{Y}$. The relation $\ll$ then satisfies (II*): (II*.1) holds by definition and (II.5).

[6] See loc. cit. [2].

[7] loc. cit. [1], chap. IX, § 1, no. 5.

(II*.2) If $A \ll B$ then, by (II.2), $X < Y_1 < X_1 < Y$ whence, by (II.5), $\overline{X} < \mathring{Y}_1 < \overline{X_1} < \mathring{Y}$, and putting $A_1 = \overline{X_1}$, $B_1 = \mathring{Y}_1$, $A \ll B_1 < A_1 \ll B$.

(II*.3) Given any open $N(x)$ there exists, by (II.3), $M(x) \prec N(x)$, whence, by (II.5); $\overline{M(x)} \ll N(x)$.

2. Conversely, it will be shown in § 3.2 and 3.1 that (II*) implies (I) and (I) implies (II).

(III) *S is uniformizable.*

Note. The following condition is ***sufficient*** for the uniformizability of S: *There exists a family $\mathfrak{F} = \{F, G, H, \ldots\}$ of sets of $S \times S$ such that*

(III.1) if D is the diagonal of $S \times S$, then $D < F$ for every $F \in \mathfrak{F}$;

(III.2) $F = F^{-1}$ for every $F \in \mathfrak{F}$;

(III.3) to any $F \in \mathfrak{F}$ there exist $G, H \in \mathfrak{F}$ such that $(G \cap H)^2 < F$;

(III.4) to any open neighbourhood $N(x)$ of any $x \in S$ there exists $F \in \mathfrak{F}$ such that $N(x) = F(x)$, and, conversely, every $F(x)$ is a neighbourhood of x.

Proof. Let $\mathfrak{U}$ be the family of all finite intersections $U = \bigcap_{i=1}^{n} F_i$ $(n \geq 1)$ of sets $F_i \in \mathfrak{F}$; $\mathfrak{U}$ is a basis of entourages of a uniform structure of S, i.e.

(α) $\mathfrak{U}$ is a filter basis,

(β) $D < U$ for every $U \in \mathfrak{U}$,

(γ) $U = U^{-1}$ for every $U \in \mathfrak{U}$,

(δ) to any $U \in \mathfrak{U}$ there exists $V \in \mathfrak{U}$ such that $V^2 < U$.

(α), (β), (γ) obviously hold in virtue of (III.1) and (III.2). (δ) follows from (III.3): if $U = \bigcap_{i=1}^{n} F_i$ and $(G_i \cap H_i)^2 < F_i$, put $V = \bigcap_{i=1}^{n} (G_i \cap H_i)$.

It remains to show that the topology defined by $\mathfrak{U}$ is identical with the given topology of S, i.e.

(ε) (i) To every neighbourhood $M(x)$ there exist $U \in \mathfrak{U}$ such that $U(x) < M(x)$,

(ii) $U(x)$ is a neighbourhood of x for every $U \in \mathfrak{U}$.

(i) follows from (III.4) as $M(x)$ contains an open neighbourhood $N(x)$, and $N(x) = F(x)$ for some $F \in \mathfrak{F} < \mathfrak{U}$; (ii) follows from (III.4) as $U = \bigcap_{i=1}^{n} F_i$ implies $U(x) = \bigcap_{i=1}^{n} F_i(x)$ and every $F_i(x)$ is a neighbourhood of x.

(IV) *There exists in S a family Δ of filters with the following properties:*

(IV.1) the neighbourhood filter φ_x of any $x \in S$ belongs to Δ;

(IV.2) if $\xi \in \Delta$, $B \in \xi$ then there exists $A \in \xi$ such that every $\eta \in \Delta$ which is compatible with A contains B;

(IV.3) every open filter of S is compatible with a filter $\xi \in \Delta$.

Note 1. The A, B of (IV.2) satisfy $A < B$ and, more precisely, $\bar{A} < \mathring{B}$.

Proof. Any neighbourhood filter φ_x compatible with A contains B; but φ_x is compatible with A if (and only if) $x \in \bar{A}$; thus B is a neighbourhood of every $x \in \bar{A}$ whence $A < \mathring{B}$.

Note 2. Condition (IV.3) may be replaced by either of these:

(IV.3′) every filter of S is compatible with a filter $\xi \in \Delta$;

(IV.3″) every ultrafilter of S contains a filter $\xi \in \Delta$.

Proof. Clearly, (IV.3″) implies (IV.3′) and (IV.3′) implies (IV.3). It remains to show that (IV.3) implies (IV.3″). Let φ be any ultrafilter, φ' the open filter generated by the open sets of φ; by (IV.3) φ' is compatible with some $\xi \in \Delta$; we assert that $\xi < \varphi$ i.e. that $B \in \xi$ implies $B \in \varphi$. If $B \in \xi$, construct $A \in \xi$ by (IV.2); then, by Note 1, $A < \bar{A} < B$, and it suffices to show that $\bar{A} \in \varphi$. Suppose $\bar{A} \notin \varphi$; then, as φ is an ultrafilter, $\mathbf{C}\bar{A} \in \varphi$; as $\mathbf{C}\bar{A}$ is open, $\mathbf{C}\bar{A} \in \varphi'$; but $\bar{A} \in \xi$, and this contradicts the compatibility of φ' and ξ.

Note 3. The following is a consequence of (IV):

(IV.4) If $\xi, \eta \in \Delta$, $\xi \neq \eta$, then ξ, η are incompatible.

Proof. Suppose ξ, η are compatible, and let B be any element of ξ. Construct $A \in \xi$ by (IV.2); as η is compatible with A, $B \in \eta$. Similarly every element of η is an element of ξ, i.e. $\xi = \eta$.

Note 4. Conditions (IV) together with the condition that S is a T_0-space are a simplified version of the conditions given by A. Gerolini[4]) for the compactifiability of S.

§ 3. Proof of the equivalence of (I) to (IV)

3.1. (I) implies (II)

Define $A < B$ if and only if there is a string with least element A and greatest element B. (II.1), (II.2), (II.4) hold in virtue of § 1 (ii)—(v), and (II.3) follows from (I).

3.2. (II) implies (I)

As (II) implies (II*) it is sufficient to show that (II*) implies (I). Given any open neighbourhood $N(x)$ of x, construct $M(x)$, by (II*.3), such that $M = M(x) < N(x) = N$. Using (II*.2), interpolate successively:

1. step: $M < N_1^1 < M_1^1 < N$,

2. step: $M < N_1^2 < M_1^2 < N_1^1 < M_1^1 < N_2^2 < M_2^2 < N$,

etc., defining, at the n-th step, M_i^n, N_i^n $(1 \leq i \leq 2^{n-1})$. By (II*.1) the family consisting of M, N and all M_i^n, N_i^n is a string of neighbourhoods of x with greatest element N.

3.3. (II) implies (III)

It is sufficient to show that (II*) implies (III). Define $\mathfrak{F}$ as the family of all F of the form

$$F = B \times B \cup \mathbf{C}A \times \mathbf{C}A \qquad \text{where } A < B.$$

We have to verify (III.1)—(III.4).

(III.1) $D < F$: as $A < B$, by (II*.1), $x \in S$ implies $x \in B$ or $x \in \mathbf{C}A$ whence $(x, x) \in F$.

(III.2) $F = F^{-1}$, as F is symmetrical by construction.

(III.3) Given $F = B \times B \cup \complement A \times \complement A$ where $A < B$, find A_1, B_1, by (II*.2) such that $A < B_1 \subset A_1 < B$ and define G, H by

$$G = B_1 \times B_1 \cup \complement A \times \complement A, \quad H = B \times B \cup \complement A_1 \times \complement A_1.$$

We assert that $(G \cap H)^2 \subset F$. Let $(x, y) \in (G \cap H)^2$. This means that z exists such that $(x, z) \in G \cap H$, $(z, y) \in G \cap H$. We have to show that $(x, y) \in F$. Suppose $(x, y) \notin F$. Then *either* $x \in B$ whence $y \notin B$, $y \in \complement A$, $x \notin \complement A$; *or* $x \in \complement A$, $y \notin \complement A$, $y \in B$, $x \notin B$. Because of the symmetry it is sufficient to consider the first case; as $x \notin \complement A$ and $(x, z) \in G$, $z \in B_1$; and as $y \notin B$ and $(z, y) \in H$, $z \in \complement A_1$; this is a contradiction, as $B_1 \cap \complement A_1 = O$.

(III.4) Given any open $N(x)$, find $M(x) < N(x)$ by (II*.3) and put $F = N \times N \cup \complement M \times \complement M$; as $x \notin \complement M$, $F(x) = N(x)$. — Conversely, if $F = B \times B \cup \complement A \times \complement A$, $A < B$, then B and $\complement A$ are open by (II*.1), and as $F(x)$ is B or $\complement A$ or S, it is a neighbourhood of x.

3.4. (III) implies (II)

Let $\mathfrak{U}$ be the filter of entourages of the given uniform structure of S and define

$A < B$ if and only if there exists $U \in \mathfrak{U}$ such that $U(A) \cap U(\complement B) = O$.

We have to verify (II.1)–(II.4).

(II.1) If $A < B$, then $A \subset U(A) \subset \complement U(\complement B) \subset \complement\complement B = B$, and as $U(A)$ is a neighbourhood of A, $A \subset \mathring{B}$.

(II.4) If $A < B$, then $U(\complement B) \cap U(\complement\complement A) = O$, i.e. $\complement B < \complement A$.

(II.2) If $A < B$, i.e. $U(A) \cap U(\complement B) = O$, choose a symmetrical V such that $V^2 \subset U$. Put $B_1 = V^2(A)$; then, by the Lemma of § 1,

$$V(\complement B_1) = V(\complement V^2(A)) \subset \complement V(A), \quad \text{i.e.} \quad V(A) \cap V(\complement B_1) = O,$$

whence $A < B_1$. — Similarly, starting with $\complement B < \complement A$ we get $\complement B < \complement A_1$ where $\complement A_1 = V^2(\complement B)$, and hence $A_1 < B$. It follows that $B_1 \subset A_1$, i.e. $B_1 \cap \complement A_1 = O$ because

$$B_1 \cap \complement A_1 = V^2(A) \cap V^2(\complement B) \subset U(A) \cap U(\complement B) = O.$$

(II.3) Let $N(x)$ be any neighbourhood of x; then $N(x) = U(x)$ for some $U \in \mathfrak{U}$; choose a symmetrical $V \in \mathfrak{U}$ such that $V^3 \subset U$, and put $M(x) = V(x)$; then, as $V^3(x) \subset U(x) = N$, $\complement N \subset \complement V^3(x)$ and

$$V(\complement N) \subset V(\complement V^3(x)) \subset \complement V^2(x) = \complement V(M),$$

by the Lemma of § 1, i.e. $V(M) \cap V(\complement N) = O$ whence $M < N$.

3.5. (III) implies (IV)

As above, let $\mathfrak{U}$ be the filter of entourages. The envelope φ^* of any ultrafilter φ of S is the filter generated by the sets $U(X)$ where $U \in \mathfrak{U}$, $X \in \varphi$. Let Δ be the family of all φ^*.

(IV.3″) is satisfied as $\varphi^* \subset \varphi$, and it remains to verify (IV.1), (IV.2).

(IV.1) For any $x \in S$ let φ be the ultrafilter generated by $\{x\}$. Every $X \in \varphi$ contains x, hence every $U(X)$ is a neighbourhood of x. On the other hand every neighbourhood $N(x)$ of x is of the form $U(x) = U(\{x\})$. Hence φ^* is identical with the neighbourhood filter φ_x of x, i.e. $\varphi_x \in \Delta$.

(IV.2) Let $\varphi^* \in \Delta$, $B \in \varphi^*$, i.e. $B > U(X)$, $U \in \mathfrak{U}$, $X \in \varphi$. Choose a symmetrical $V \in \mathfrak{U}$ such that $V^3 < U$ and let $A = V(X)$; then $A \in \varphi^*$. Let $\psi^* \in \Delta$ be compatible with $A = V(X)$; we assert that $V^2(X) \in \psi$ whence $V^3(X) \in \psi^*$ and, as $V^3(X) < U(X) < B$, $B \in \psi^*$. Suppose $V^2(X) \notin \psi$; then $\complement V^2(X) \in \psi$, whence $V(\complement V^2(X)) \in \psi^*$; by the Lemma of § 1, $V(\complement V^2(X)) < \complement V(X)$, whence $\complement V(X) = \complement A \in \psi^*$ which contradicts the compatibility of ψ^* with A.

3.6. (IV) implies (III)

Let Δ be a family of filters satisfying (IV). A family $\{A_1, \ldots, A_n\}$ of a finite number $n \geq 1$ of subsets of S will be called a *special finite covering* (s.f.c.) of S if $\bigcup_{i=1}^{n} A_i = S$ and if every $\xi \in \Delta$ has at least one A_i as element. Let $\mathfrak{U}$ be the family of all sets $U < S \times S$ of the form $U = \bigcup_{i=1}^{n} A_i \times A_i$ where $\{A_1, \ldots, A_n\}$ is any s.f.c. We verify the conditions (α)—(ε) as in § 2 (III).

(α) $\mathfrak{U}$ is a filter basis: $\{S\}$ is a s.f.c., hence $S \times S \in \mathfrak{U}$, i.e. $\mathfrak{U}$ is not empty; by definition $\mathfrak{U}$ does not contain the null-set O. If $U = \bigcup_{i=1}^{n} A_i \times A_i \in \mathfrak{U}$, $V = \bigcup_{j=1}^{m} B_j \times B_j \in \mathfrak{U}$ put $C_{ij} = A_i \cap B_j$; then $\bigcup_{i,j} C_{ij} = S$; and as every $\xi \in \Delta$ contains an A_i and a B_j, it contains C_{ij} so that $\{C_{ij}\}$ is a s.f.c.; $W = \bigcup_{i,j} C_{ij} \times C_{ij}$ belongs to $\mathfrak{U}$ and satisfies $W < U \cap V$.

(β) $D < U$ for every $U \in \mathfrak{U}$, by definition.

(γ) $U = U^{-1}$ for every $U \in \mathfrak{U}$, by definition.

(δ) To any $U \in \mathfrak{U}$ there exists $V \in \mathfrak{U}$ such that $V^2 < U$.

Proof. Suppose there exists $U = \bigcup_{i=1}^{n} A_i \times A_i \in \mathfrak{U}$ such that for *no* $V \in \mathfrak{U}$, $V^2 < U$. Then for every $V \in \mathfrak{U}$, V^2 is compatible with $\complement U$. The intersections $V^2 \cap \complement U$ form a filterbasis in $S \times S$ generating a filter $\mathfrak{F}$, and $\mathfrak{F}$ is contained in an ultrafilter $\mathfrak{G}$. Clearly $\complement U \in \mathfrak{G}$ and $V^2 \in \mathfrak{G}$ for all $V \in \mathfrak{U}$.

The projections φ, ψ of $\mathfrak{G}$ (projecting $S \times S$ on the first and the second factor S) are ultrafilters of S. By (IV.3″) there exist $\xi, \eta \in \Delta$ such that $\xi < \varphi$ and $\eta < \psi$. We note that $P \in \xi$, $Q \in \eta$ implies $P \in \varphi$, $Q \in \psi$ and hence $P \times Q \in \mathfrak{G}$.

By definition, one $A_i \in \xi$. $A_i \notin \eta$, because $A_i \in \eta$ would imply $A_i \times A_i \in \mathfrak{G}$ whence $U \in \mathfrak{G}$, contradicting $\complement U \in \mathfrak{G}$. Hence $\xi \neq \eta$ and, by (IV.4), ξ, η are incompatible, i.e. there exist $X_1 \in \xi$, $X_2 \in \eta$ with $X_1 \cap X_2 = O$.

By (IV.2) and (IV) Note 1, there exist $Y_1 \in \xi$, $Y_2 \in \eta$, $Y_1 < X_1$, $Y_2 < X_2$ such that every $\zeta \in \Delta$ compatible with Y_1 resp. Y_2 contains X_1 resp. X_2. Put $X_3 = \complement(Y_1 \cap Y_2)$.

$\{X_1, X_2, X_3\}$ is a s.f.c.; for $\bigcup_{i=1}^{3} X_i > Y_1 \cup Y_2 \cup \complement(Y_1 \cup Y_2) = S$; and if $\zeta \in \Delta$, then either $X_1 \in \zeta$ or $X_2 \in \zeta$ or ζ is incompatible with Y_1 and Y_2; in the last

case there exist $Z_1, Z_2 \in \zeta$ such that $Z_1 \cap Y_1 = Z_2 \cap Y_2 = O$, whence, putting $Z = Z_1 \cap Z_2$, $Z \in \zeta$, $Z \cap (Y_1 \cup Y_2) = O$, i.e. $Z \subset \mathbf{C}(Y_1 \cup Y_2) = X_3$ and $X_3 \in \zeta$. Hence $V = \bigcup_{i=1}^{3} X_i \times X_i \in \mathfrak{U}$.

As $V^2 \in \mathfrak{G}$ and $Y_1 \times Y_2 \in \mathfrak{G}$, $V^2 \cap (Y_1 \times Y_2) \neq O$. Let $(y_1, y_2) \in V^2 \cap (Y_1 \times Y_2)$; then $(y_1, y_2) \in Y_1 \times Y_2$, i.e. $y_1 \in Y_1$, $y_2 \in Y_2$, and $(y_1, y_2) \in V^2$, i.e. z exists such that $(y_1, z) \in V$, $(z, y_2) \in V$, which means that for some j, k

$$(y_1, z) \in X_j \times X_j, \quad (z, y_2) \in X_k \times X_k.$$

As $y_1 \in Y_1$, $y_1 \notin \mathbf{C}(Y_1 \cup Y_2) = X_3$, and as $Y_1 \subset X_1$ and $X_1 \cap X_2 = O$, $y_1 \notin X_2$; hence $j = 1$, and similarly $k = 2$. But this gives the contradiction $z \in X_1$, $z \in X_2$.

(ε) (i) Let $N(x) = X_1$ be any neighbourhood of x.[8] By (IV.1), (IV.2) and (IV) Note 1, there exists a neighbourhood $M(x) \subset N(x)$ such that every $\zeta \in \Delta$ compatible with M contains N; put $X_2 = \mathbf{C}M$. $\{X_1, X_2\}$ is a s.f.c.; for $X_1 \cup X_2 = N \cup \mathbf{C}M = S$: and, for any $\zeta \in \Delta$, either $X_1 \in \zeta$ or ζ is incompatible with M; in the last case there exists $Z \in \zeta$ such that $Z \cap M = O$, i.e. $Z \subset \mathbf{C}M = X_2$ whence $X_2 \in \zeta$. Taking $U = X_1 \times X_1 \cup X_2 \times X_2$ we have $U \in \mathfrak{U}$ and $N(x) = U(x)$.

(ii) Let $U = \bigcup_{i=1}^{n} A_i \times A_i \in \mathfrak{U}$, $x \in S$; as the neighbourhood filter φ_x of x belongs to Δ, one $A_i \in \varphi_x$; hence A_i is a neighbourhood of x, and as $A_i \subset U(x)$, $U(x)$ is a neighbourhood of x.

Note. As the entourages $U \in \mathfrak{U}$ are defined by finite converings of S, the uniform structure is totally bounded.

§ 4. Compactifiability

As pointed out in the introduction, a topological space S can be embedded in a compact (= bicompact) space if and only if S is a T_0-space satisfying one of the conditions (I)—(IV). We sketch a proof following A. Gerolini[4]).

1. Let S be a subspace of the compact space T; it may be assumed that S is dense in T. S is a T_0-space. We show that S satisfies (IV) where Δ is the family of the traces ψ'_u on S of the neighbourhood filters ψ_u of the points $u \in T$.

(IV.1) If φ_x is the neighbourhood filter in S of any $x \in S$ then $\varphi_x = \psi'_x \in \Delta$.

(IV.2) Let $\xi \in \Delta$, $B \in \xi$; then $\xi = \psi'_u$ for some $u \in T$ and $B = R \cap S$ for some $R \in \psi_u$; as T is regular, there exist an open neighbourhood Q and a closed neighbourhood P of u such that $P \subset Q \subset R$; put $A = P \cap S$. Let $\eta \in \Delta$ be compatible with A; $\eta = \psi'_v$ for some $v \in T$. ψ_v is compatible with P (if $V \in \psi_v$ then $V \cap S \in \psi'_v$, $(V \cap S) \cap A \neq O$, whence $V \cap P \neq O$). This means that v adheres to P, and as P is closed $v \in P$; as Q is open and $P \subset Q \subset R$, R is a neighbourhood of v, i.e. $R \in \psi_v$ and $B = R \cap S \in \psi'_v = \eta$.

(IV.3′) Let φ be any filter of S. As T is compact, φ, considered as filter basis in T, has an adherent point $u \in T$; this means that φ is compatible with ψ_u and hence with $\psi'_u \in \Delta$.

2. Conversely, let S be a T_0-space with a family Δ of filters satisfying (IV).

Δ can be topologized by taking as a basis of open sets the family of all sets $\Omega_B \subset \Delta$ consisting of all $\xi \in \Delta$ with $B \in \xi$ where B is any open set of S. (A basis of closed sets is then given by the family of all sets $\Gamma_A \subset \Delta$ consisting of all $\eta \in \Delta$ compatible with A where A is any closed set of S.) Using (IV.4) and (IV.2), Δ is proved to be (i) a T_2-space, (ii) a regular space.

As S is a T_0-space, it follows, using (IV.1), that the map $x \to \varphi_x$ is a one-to-one map of S onto a subspace Σ of Δ, and it can easily be shown that the map is a homeomorphism of S on Σ, and that Σ is dense in Δ.

The proof is completed by showing that Δ is compact. Let $\mathfrak{F}$ be any filter of Δ, $\mathfrak{F}'$ the filter generated by the open sets of $\mathfrak{F}$, $\mathfrak{F}''$ the trace of $\mathfrak{F}'$ on Σ. By the homeomorphism of Σ and S, $\mathfrak{F}''$ corresponds to an open filter φ of S which, by (IV.3), is compatible with a filter $\xi \in \Delta$. It is easily seen that ξ, considered as point of Δ, adheres to $\mathfrak{F}'$, and it follows from the regularity of Δ, that ξ adheres to $\mathfrak{F}$.

Note. Fan and Gottesman[5]) proved that a topological space S is compactifiable if S is regular and possesses a normal base, i.e. a base $\mathfrak{B}$ of open sets satisfying

(1) if A, $B \in \mathfrak{B}$ then $A \cap B \in \mathfrak{B}$,

(2) if $A \in \mathfrak{B}$ then $\complement \bar{A} \in \mathfrak{B}$,

(3) to any open $U \subset S$ and any $A \in \mathfrak{B}$ such that $\bar{A} \subset U$ there exists $B \in \mathfrak{B}$ such that $\bar{A} \subset B \subset \bar{B} \subset U$.

We show that conditions (1) and (2) are not required. Let S be a regular space with a base $\mathfrak{B}$ of open sets satisfying (3). Define

$$X < Y \text{ if and only if } Y \text{ is open, } X = \bar{A} \text{ for some } A \in \mathfrak{B}, \text{ and } X \subset Y.$$

As S is a T_0-space, it is sufficient to verify that the relation $<$ satisfies (II*). (II*.1) holds by definition.

(II*.2) If $X = \bar{A} < Y$, then there exists by (3) $B \in \mathfrak{B}$ such that

$$X = \bar{A} \subset B \subset \bar{B} \subset Y \quad \text{whence} \quad X < B < \bar{B} < Y.$$

(II*.3) Let $N(x)$ be an open neighbourhood of x. As S is regular, there exists a neighbourhood $M(x)$ such that $\overline{M(x)} \subset N(x)$; $M(x)$ contains a basic open neighbourhood $L(x) \in \mathfrak{B}$; as $\overline{L(x)} \subset N(x)$ we have $\overline{L(x)} < N(x)$.

Department of Mathematics, University of Melbourne

(Eingegangen am 3. September 1956)

JURKAT, W. B.
Math. Zeitschr. Bd. 67, S. 211—222 (1957)

Über die Umkehrung des Abelschen Stetigkeitssatzes mit funktionentheoretischen Methoden

Von

W. B. JURKAT

Diese Note befaßt sich mit der Verfeinerung und der systematischen Untersuchung der in [6] aufgestellten Beweismethoden. Ausgangspunkt ist der Abelsche Stetigkeitssatz in der allgemeinen Form: Wenn eine (reelle oder komplexe) Reihe $\sum a_\nu$ gegen Null konvergiert oder allgemeiner für ein $k > -1$

$$(\mathrm{C,k}) \qquad S_n^{k+1}[a_\nu] = \sum_{\nu=0}^{n} \binom{n-\nu+k}{n-\nu} a_\nu = o(n^k) \quad \text{für} \quad n \to \infty$$

ist, so gilt

$$(\mathrm{A}) \qquad \begin{cases} f(z) = \sum_{\nu=0}^{\infty} a_\nu z^\nu & \text{konvergiert für } |z|<1 \text{ und} \\ f(z) = o(1) & \text{für reelle } z \to 1-0. \end{cases}$$

Es gilt sogar $f(z) = o(1)$ für $z \to 1$ in jedem Winkelraum $|1-z| < M(1-|z|)$ mit $M > 1$. Die wichtigsten Umkehrungen rühren von LITTLEWOOD her [vgl. 2, Chap. VII]:

Satz I. (A) $\Rightarrow$ (C, 1), *falls* $s_n = S_n^1[a_\nu] = O(1)$.

Satz II. (A) $\Rightarrow$ (C, 0), *falls* $n a_n = O(1)$.

Jeder der Sätze kann aus dem anderen hergeleitet werden. Die zusätzlich auftretende Umkehrbedingung läßt sich erheblich verallgemeinern: Zum Beispiel genügt bei Satz II die Bedingung

$$(\mathrm{T}) \qquad s_{n'} - s_n = \sum_{n+1}^{n'} a_\nu = o(1) \quad \text{für} \quad 1 < \frac{n'}{n} \to 1,$$

die zugleich notwendig für Konvergenz ist.

Will man Satz II mit der Bedingung (T) oder in noch allgemeinerer Form beweisen, so zeigt man im allgemeinen erst Beschränktheit, dann C-Limitierbarkeit (Satz I) und schließlich Konvergenz von s_n. Damit wird die Hauptschwierigkeit auf Satz I hinausgespielt, in den die allgemeine Umkehrbedingung gar nicht mehr eingeht. Der bisher einfachste Beweis für Satz I wurde von KARAMATA gegeben, und die Beweismethode ist unlängst von Wielandt so variiert worden, daß sie auch unmittelbar auf Satz II anwendbar wird [11], [5]. Die Methode kann durch das Stichwort Approximation gekennzeichnet werden.

Demgegenüber wurde in [6] ein neuer einfacher Beweis von Satz II und seinen Verallgemeinerungen angegeben, der ähnlich wie oben in drei Schritten vorgeht und sich vor allem auf die Cauchysche Koeffizientenformel stützt. Solange man allgemeine Umkehrbedingungen im Auge hat, erscheint eine solche Dreiteilung des Problems völlig angemessen. Es erhebt sich aber die Frage, ob sich der funktionentheoretische Beweis des speziellen Satzes II nicht ähnlich vereinfachen läßt wie das bei der Karamataschen Beweismethode der Fall war. Tatsächlich werden durch eine Variation der bisherigen funktionentheoretischen Methode die Sätze I und II einer direkten Behandlung zugänglich (§§ 2 und 3).

Der entscheidende Gedanke besteht in der Einführung allgemeinerer Umkehrbedingungen, die sich besser funktionentheoretisch erfassen lassen. Für jede solche Umkehrbedingung können alle Indizes k mit der Eigenschaft $(A) \Rightarrow (C, k)$ bestimmt werden. Es zeigt sich, daß zu jeder „arithmetischen" Umkehrbedingung eine „funktionentheoretische" Umkehrbedingung vorhanden ist, die die gleiche „Umkehrwirkung" hat. So entsteht eine Vielfalt von Umkehrsätzen, deren Gesamtheit vollständig aufgestellt wird (§ 7). Besonderes Interesse verdienen die extremen Umkehrbedingungen, unter die sich auch die früheren Ergebnisse einordnen lassen (§§ 5 und 6). Die neuen Methoden führen auch zu weiteren Anwendungen beim Borelschen Verfahren, auf die wir bei anderer Gelegenheit eingehen werden [vgl. 7].

1. Der Ansatzpunkt aller folgenden Beweise besteht in der funktionentheoretischen Formulierung der Umkehrbedingung. Bei Satz I handelt es sich um die Frage, was $s_n = O(1)$ für die Funktion

$$F(z) = \sum_{\nu=0}^{\infty} s_\nu z^\nu \quad \text{mit} \quad z = r e^{i\varphi}, \quad 0 < r < 1, \ |\varphi| \leq \pi \tag{1}$$

bedeutet. Offenbar folgt

$$F(z) = O(1)\,(1-r)^{-1}, \tag{2}$$

aber dies folgt auch schon aus $\sum |s_\nu| r^\nu = O(1)\,(1-r)^{-1}$ und damit aus der schwächeren Bedingung

$$\sum_{\nu=0}^{n} |s_\nu| = O(n). \tag{3}$$

Übrigens ist allgemein für jedes $\gamma > 0$

$$\sum_{\nu=0}^{n} |s_\nu| = O(n^\gamma) \text{ gleichbedeutend mit } \sum_{\nu=0}^{\infty} |s_\nu| r^\nu = O(1)\,(1-r)^{-\gamma}, \tag{4}$$

wie man durch eine einfache Überlegung erkennt $\left(\text{setze } r = 1 - \frac{1}{n}\right)$. Mit der Umkehrbedingung (3) statt $s_n = O(1)$ ist aber Satz I nicht mehr richtig, wie wir noch zeigen werden. Daher kommt es darauf an, aus $s_n = O(1)$ mehr herauszuholen wie z. B.

$$\sum_{\nu=0}^{n} |s_\nu|^2 = O(n). \tag{5}$$

Dies ist nach (4) äquivalent mit $\sum |s_\nu|^2 r^\nu = O(1)(1-r)^{-1}$ und daher nach der Parsevalschen Gleichung mit

$$(6) \qquad \int_{-\pi}^{\pi} |F(re^{i\varphi})|^2 d\varphi = 2\pi \sum_{\nu=0}^{\infty} |s_\nu|^2 r^{2\nu} = O\left(\frac{1}{1-r}\right).$$

Die Bedingung (5) läßt sich also vollständig funktionentheoretisch ausdrücken[1]). Da (5) mehr verlangt als (3), so bedeutet auch (6) mehr als (2). Mit der Umkehrbedingung (6) bzw. (5) statt $s_n = O(1)$ läßt sich nun Satz I bereits beweisen, wie wir noch zeigen werden.

Es zeigt sich also, daß die Umkehrbedingung $s_n = O(1)$ zweckmäßig durch Bedingungen der Form

$$(7) \qquad \sum_{\nu=0}^{n} |s_\nu|^\varrho = O(n) \quad \text{mit} \quad \varrho \geq 1$$

zu ersetzen ist, die besser funktionentheoretisch ausgedrückt werden können. Diese Bedingungen verlangen um so mehr, je größer ϱ ist, und gehen für $\varrho \to +\infty$ in die Bedingung $s_n = O(1)$ über. Nach der Ungleichung von HAUSDORFF-YOUNG gilt für $1 < \varrho < +\infty$, $\frac{1}{\varrho} + \frac{1}{\varrho'} = 1$:

$$\int_{-\pi}^{\pi} |F(z)|^{\varrho'} d\varphi = O(1) \left(\sum_{\nu=0}^{\infty} |s_\nu r^\nu|^\varrho \right)^{\varrho'/\varrho} \quad \text{falls } \varrho \leq 2$$

bzw.

$$\left(\sum_{\nu=0}^{\infty} |s_\nu r^\nu|^\varrho \right)^{\varrho/\varrho'} = O(1) \int_{-\pi}^{\pi} |F(z)|^{\varrho'} d\varphi \quad \text{falls } \varrho \geq 2.$$

Daher entspricht der arithmetischen Bedingung (7) die funktionentheoretische Bedingung

$$(7') \qquad \int_{-\pi}^{\pi} |F(z)|^{\varrho'} d\varphi = O(1)/(1-r)^{\varrho'-1},$$

die für $\varrho \leq 2$ schwächer und für $\varrho \geq 2$ stärker als (7) ist [im weiteren Sinn][2]).

Was hier über die Umkehrbedingung $s_n = O(1)$ gesagt wurde, gilt entsprechend für $n a_n = O(1)$, wobei man statt $F(z)$ die Funktion $z f'(z) = \sum_{\nu=0}^{\infty} \nu a_\nu z^\nu$ zu betrachten hat. Wegen $(1-z)F(z) = f(z)$ lassen sich alle Umkehrbedingungen sowohl mit $F(z)$ als auch mit $f(z)$ ausdrücken. Will man dabei z durch r und φ ausdrücken, so ist die Abschätzung

$$(8) \qquad K_1 |1-z| \leq \operatorname{Max}(1-r, |\varphi|) \leq K_2 |1-z|$$

mit positiven Konstanten $K_{1,2}$ nützlich.

[1]) Für die Bedeutung der Parsevalschen Gleichung bei genauen Integralabschätzungen vgl. auch [8, Kap. I].

[2]) Nur für Vergleiche dieser Art wird im folgenden die Ungleichung von HAUSDORFF-YOUNG benutzt werden. Die Beweise der Sätze sind davon unabhängig.

2. Statt Satz I beweisen wir allgemeiner den folgenden

Satz 1. *Unter der Umkehrbedingung*

$$\text{(9)} \qquad \sum_{\nu=0}^{n} |s_\nu|^2 = O(n) \quad \text{oder} \quad \int_{-\pi}^{\pi} |F(z)|^2 d\varphi = O\left(\frac{1}{1-r}\right)$$

gilt $(A) \Rightarrow (C, k)$ *für* $k > \frac{1}{2}$; *aber nicht für* $k = \frac{1}{2}$.

Beweis. Wir wissen schon, daß die beiden Bedingungen von (9) äquivalent sind und $F(z) = O(1)(1-r)^{-1}$ zur Folge haben. Also ist $f(z) = (1-z)F(z) = O(1)$ in jedem Winkelraum und $= o(1)$ für reelle $z \to 1-0$, so daß nach einem Satz von MONTEL [**10**, S. 170/71]

$$\text{(10)} \qquad f(z) = o(1) \qquad \text{für } z \to 1 \text{ in jedem Winkelraum}$$

gilt. Aus $G(z) = (1-z)^{-k}F(z) = (1-z)^{-k-1}f(z) = \sum_n S_n^{k+1}[a_\nu] z^n$ ergibt sich nach der CAUCHYschen Koeffizientenformel

$$\text{(11)} \qquad S_n^{k+1}[a_\nu] = \frac{1}{2\pi}\left(\int_{|\varphi| \le m/n} + \int_{m/n \le |\varphi| \le \pi}\right) \frac{G(z)}{z^n}\, d\varphi = A + B,$$

wobei $r = 1 - \frac{1}{n}$ und $n \gtrsim \frac{n}{m} \gtrsim 2$ sein soll. Nach der Ungleichung von CAUCHY-SCHWARTZ und (8) gilt gleichmäßig in n, m

$$B = O(1)\left\{\int_{-\pi}^{\pi} |F(z)|^2 d\varphi\right\}^{\frac{1}{2}} \left\{\int_{m/n}^{\pi} \frac{d\varphi}{\varphi^{2k}}\right\}^{\frac{1}{2}} = O(n^{\frac{1}{2}}) \cdot \left(\frac{n}{m}\right)^{k-\frac{1}{2}}.$$

Also kann man zu $\varepsilon > 0$ ein $m \gtrsim 2\pi$ so finden, daß für alle großen n durchweg $|B| \le \frac{\varepsilon}{2} n^k$ ist. Wird m in dieser Weise fest gewählt, so liegen die in A vorkommenden z ganz in einem festen Winkelraum, und wegen (10) und (8) gilt

$$A = o(1) \int_{|\varphi| \le m/n} \frac{d\varphi}{(1-r)^{k+1}} = o(n^k) \qquad \text{für } n \to \infty.$$

Daher ist insgesamt $|S_n^{k+1}[a_\nu]| \le \varepsilon n^k$ für alle großen n. Dies beweist den Satz für $k > \frac{1}{2}$, während man für $k = \frac{1}{2}$ die Folge

$$s_n = \begin{cases} \sqrt{n} & \text{für} \quad n = 2^m \quad (m = 0, 1, \ldots) \\ 0 & \text{sonst} \end{cases}$$

betrachtet. Für sie gilt $\sum_0^n |s_\nu|^2 = O(n)$, $\sum_0^n s_\nu = O(\sqrt{n}) = o(n)$ und folglich (A), aber $s_n \neq o(\sqrt{n})$ und folglich nicht $(C, \frac{1}{2})$.

3. Statt Satz II beweisen wir allgemeiner den folgenden

Satz 2. *Unter der Umkehrbedingung*

$$\text{(12)} \qquad \sum_{\nu=0}^{n} |\nu a_\nu|^2 = O(n) \quad \text{oder} \quad \int_{-\pi}^{\pi} |z f'(z)|^2 d\varphi = O\left(\frac{1}{1-r}\right)$$

gilt $(A) \Rightarrow (C, k)$ *für* $k > -\frac{1}{2}$, *aber nicht für* $k = -\frac{1}{2}$.

Beweis. Wir wissen schon, daß die beiden Bedingungen von (12) äquivalent sind. Ferner folgt nach der Ungleichung von CAUCHY-SCHWARTZ und (8)

$$(13)\qquad \begin{cases} f(z) = f(r) + \int\limits_0^{\varphi} f'(r e^{it})\, r e^{it} i\, dt = O(1) + O(1)\,|\varphi|^{\frac{1}{2}} (1-r)^{-\frac{1}{2}} \\ \quad = O(1)\left(\dfrac{|1-z|}{1-|z|}\right)^{\frac{1}{2}}, \text{ also } = O(1) \text{ in jedem Winkelraum.} \end{cases}$$

Wie beim Beweis von Satz 1 ergibt sich wieder (10) und (11), wobei aus B durch partielle Integration $B_1 + B_2 + B_3$ entsteht:

$$B = \frac{1}{2\pi} \cdot \frac{z^{-n}}{-in} \cdot \frac{f(z)}{(1-z)^{k+1}} \Bigg|_{\varphi = -\frac{m}{n}}^{\varphi = \frac{m}{n}} + \frac{1}{2\pi} \int\limits_{m/n \le |\varphi| \le \pi} \frac{z^{-n}}{in} \left\{ \frac{i z f'(z)}{(1-z)^{k+1}} + \frac{(k+1)\, i z f(z)}{(1-z)^{k+2}} \right\} d\varphi .$$

Nach CAUCHY-SCHWARTZ und (8), (13) gilt gleichmäßig in n, m

$$B_2 = O\left(\frac{1}{n}\right) \left\{ \int\limits_{-\pi}^{\pi} |z f'(z)|^2\, d\varphi \right\}^{\frac{1}{2}} \left\{ \int\limits_{m/n}^{\pi} \frac{d\varphi}{\varphi^{2(k+1)}} \right\}^{\frac{1}{2}} = O(n^{-\frac{1}{2}}) \left(\frac{n}{m}\right)^{k+\frac{1}{2}},$$

$$B_3 = O\left(\frac{1}{n}\right) \int\limits_{m/n}^{\pi} n^{\frac{1}{2}} \frac{d\varphi}{\varphi^{k+\frac{1}{2}}} = O(n^{-\frac{1}{2}}) \left(\frac{n}{m}\right)^{k+\frac{1}{2}}.$$

Also kann man zu $\varepsilon > 0$ ein $m \geq 2\pi$ so finden, daß für alle großen n durchweg $|B_2| \leq \frac{\varepsilon}{4} n^k$, $|B_3| \leq \frac{\varepsilon}{4} n^k$ ist. Wird m in dieser Weise festgewählt, so folgen wegen (10) und (8) wie bei Satz 1 die Abschätzungen $A = o(n^k)$, $B_1 = o(n^k)$ für $n \to \infty$. Daher ist insgesamt $|S_n^{k+1}[a_\nu]| \leq \varepsilon n^k$ für alle großen n. Dies beweist den Satz für $k > -\frac{1}{2}$, während man für $k = -\frac{1}{2}$ die Reihe mit den Gliedern

$$a_n = \begin{cases} n^{-\frac{1}{2}} & \text{für} \quad n = 2^m \quad (m = 0, 1, \ldots) \\ 0 & \text{sonst für } n \geq 1 \end{cases}$$

betrachtet, wobei a_0 aus $\sum\limits_0^{\infty} a_n = 0$ bestimmt wird. In diesem Fall gilt (12) und (A), aber $a_n \neq o(n^{-\frac{1}{2}})$ und folglich nicht $(C, -\frac{1}{2})$.

4. Die Umkehrbedingungen der Sätze I und II lassen sich in natürlicher Weise verallgemeinern zu $S_n^{\alpha}[s_\nu] = O(n^{\alpha})$ bzw. $S_n^{\alpha}[\nu a_\nu] = O(n^{\alpha})$. Entsprechende Verallgemeinerungen sind bei den Sätzen 1 und 2 möglich:

Satz 3. *Unter der Umkehrbedingung*

$$(14)\qquad \sum_{\nu=0}^{n} |S_\nu^{\alpha}[s_\mu]|^2 = O(n^{2\alpha+1}) \qquad \textit{oder} \qquad \int\limits_{-\pi}^{\pi} \left| \frac{F(z)}{(1-z)^{\alpha}} \right|^2 d\varphi = O(1)/(1-r)^{2\alpha+1}$$

mit $\alpha \geq -\frac{1}{2}$ *gilt* (A) $\Rightarrow$ (C, k) *für* $k > \alpha + \frac{1}{2}$, *aber nicht für* $k = \alpha + \frac{1}{2}$.

Satz 4. *Unter der Umkehrbedingung*

$$(15)\quad \sum_{\nu=0}^{n} |S_\nu^\alpha[\mu a_\mu]|^2 = O(n^{2\alpha+1}) \quad \textit{oder} \quad \int_{-\pi}^{\pi} \left|\frac{z f'(z)}{(1-z)^\alpha}\right|^2 d\varphi = O(1)/(1-r)^{2\alpha+1}$$

mit $\alpha > -\frac{1}{2}$ *gilt* $(A) \Rightarrow (C, k)$ *für* $k > \alpha - \frac{1}{2}$, *aber nicht für* $k = \alpha - \frac{1}{2}$.

Die Beweise verlaufen genau so wie in den Spezialfällen $\alpha = 0$: Aus (14) folgt nach CAUCHY-SCHWARTZ und (4)

$$(16)\quad \sum_{\nu=0}^{n} |S_\nu^\alpha[s_\mu]| = O(n^{\alpha+1}), \quad \text{also} \quad (1-z)^{-\alpha} F(z) = O(1)\,(1-r)^{-\alpha-1}.$$

Bei der Abschätzung von B ist die Integralform von (14) zu benutzen. Das Gegenbeispiel für Satz 3 berechnet sich aus

$$S_n^\alpha[s_\nu] = \begin{cases} n^{\alpha+\frac{1}{2}} & \text{für} \quad n = 2^m \quad (m = 0, 1, \ldots) \\ 0 & \text{sonst}. \end{cases}$$

Aus (15) und (A) folgt nach CAUCHY-SCHWARTZ und (8)

$$(17)\quad f(z) = O(1) + O(1)\,(1-r)^{-\alpha-\frac{1}{2}} \left\{\int_0^{|\varphi|} (\operatorname{Max}(t, 1-r))^{2\alpha}\, dt\right\}^{\frac{1}{2}} = O(1)\left(\frac{|1-z|}{1-|z|}\right)^{\alpha+\frac{1}{2}}.$$

Bei der Abschätzung von $B_{2,3}$ ist (15) und (17) zu verwenden. Das Gegenbeispiel für Satz 4 berechnet sich aus

$$S_n^\alpha[\nu a_\nu] = \begin{cases} n^{\alpha+\frac{1}{2}} & \text{für} \quad n = 2^m \quad (m = 0, 1, \ldots) \\ 0 & \text{sonst} \end{cases}$$

und der Bedingung $a_0 + \sum_{n=1}^{\infty} S_n^\alpha[\nu a_\nu] \Big/ \dot{n}\binom{n+\alpha}{n} = 0$ für a_0. In diesem Fall gilt $S_n^{\alpha+\frac{1}{2}}[\nu a_\nu] \neq o(n^{\alpha+\frac{1}{2}})$, also nicht $(C, \alpha - \frac{1}{2})$.

5. Wir haben bisher nur Umkehrbedingungen mit dem Exponenten $\varrho = 2$ untersucht. Nach § 1 bietet sich auch der Fall $\varrho = 1$ an, der auf die vorangehende Arbeit [6] zurückführt. Die Rechnungen sind hierbei etwas einfacher, doch werden die Sätze I und II nicht mehr unmittelbar erhalten.

Satz 5. *Unter der Umkehrbedingung*

$$(18)\quad \sum_{\nu=1}^{n} |s_\nu| = O(n) \quad \textit{oder} \quad F(z) = O\left(\frac{1}{1-|z|}\right)$$

gilt $(A) \Rightarrow (C, k)$ *für* $k > 1$, *aber nicht für* $k = 1$.

Satz 6. *Unter der Umkehrbedingung*

$$(19)\quad \sum_{\nu=0}^{n} |S_\nu^\alpha[s_\mu]| = O(n^{\alpha+1}) \quad \textit{oder} \quad (1-z)\,F(z) = O(1)\left(\frac{|1-z|}{1-|z|}\right)^{\alpha+1}$$

mit $\alpha > -1$ *gilt* $(A) \Rightarrow (C, k)$ *für* $k > \alpha + 1$, *aber nicht für* $k = \alpha + 1$.

Hierbei ist der folgende Grenzfall bemerkenswert:

Es gilt $(A) \Rightarrow (C, k)$ *für* $k > 0$, *falls* $(1-z)F(z) = O(1)$ *für* $|z| < 1$.

Beweis. Für $\alpha = 0$ geht Satz 6 in Satz 5 über. Wie bei (16) erkennt man, daß die arithmetischen Umkehrbedingungen die entsprechenden funktionentheoretischen zur Folge haben. Mit der letzteren Umkehrbedingung ist der positive Teil beider Sätze in [6] bewiesen worden. Das Gegenbeispiel berechnet sich aus

$$\sum_{\nu=0}^{n} S_\nu^\alpha [s_\mu] = \begin{cases} n^{\alpha+1} & \text{für } n = 2^m \quad (m = 0, 1, \ldots) \\ 0 & \text{sonst.} \end{cases}$$

Satz 7. *Unter der Umkehrbedingung*

$$\sum_{\nu=0}^{n} |\nu a_\nu| = O(n) \quad \text{oder} \quad z f'(z) = O\left(\frac{1}{1-|z|}\right) \tag{20}$$

gilt $(A) \Rightarrow (C, k)$ *für* $k > 0$, *aber nicht für* $k = 0$.

Satz 8. *Unter der Umkehrbedingung*

$$\sum_{\nu=0}^{n} |S_\nu^\alpha [\mu a_\mu]| = O(n^{\alpha+1}) \quad \text{oder} \quad (1-z) z f'(z) = O(1) \left(\frac{|1-z|}{1-|z|}\right)^{\alpha+1} \tag{21}$$

mit $\alpha > -1$ *gilt* $(A) \Rightarrow (C, k)$ *für* $k > \alpha$, *aber nicht für* $k = \alpha$.

Hierbei ist der folgende Grenzfall bemerkenswert:

Es gilt $(A) \Rightarrow (C, k)$ *für* $k > -1$, *falls* $(1-z) f'(z) = O(1)$ *für* $|z| < 1$.

Beweis. Für $\alpha = 0$ geht Satz 8 in Satz 7 über. Für den positiven Teil der Behauptungen genügt es, Satz 8 mit der funktionentheoretischen Umkehrbedingung zu beweisen. Wie bei (13) und (17) folgt

$$f(z) = O(1) + O(1)(1-r)^{-\alpha-1} \int_0^{|\varphi|} (\mathrm{Max}(t, 1-r))^\alpha dt = O(1) \left(\frac{|1-z|}{1-|z|}\right)^{\alpha+1}. \tag{22}$$

Bei der Abschätzung von $B_{2,3}$ ist (21) und (22) zu verwenden. Das Gegenbeispiel berechnet sich aus

$$\sum_{\nu=0}^{n} S_\nu^\alpha [\mu a_\mu] = \begin{cases} n^{\alpha+1} & \text{für } n = 2^m \quad (m = 0, 1, \ldots) \\ 0 & \text{sonst} \end{cases}$$

und der Bedingung $a_0 + \sum_{n=1}^{\infty} S_n^{\alpha+1} [\nu a_\nu] \Big/ n \binom{n+\alpha+1}{n} = 0$ für a_0.

6. Wir wollen jetzt den Fall $\varrho = +\infty$ näher untersuchen, bei dem die arithmetischen Umkehrbedingungen in den Vordergrund treten. Wir benutzen dabei den bekannten

Konvexitätssatz. *Ist* $-1 \leq \alpha < \beta$ *und*

$$S_n^k [s_\nu] = S_n^{k+1} [a_\nu] = O(n^k) \quad \text{für } k > \alpha \tag{23}$$

so folgt aus (C, k) *für* $k>\beta$ *sogar* (C, k) *für* $k>\alpha$. *Dasselbe gilt bei Ersetzung von* (23) *durch*

$$S_n^{k+1}[v\,a_v] = O(n^{k+1}) \quad \text{für} \quad k>\alpha. \tag{24}$$

Ein Beweis für den ersten Teil der Behauptung findet sich z. B. in [6], während der zweite Teil mittels der Identität

$$(n+k+1)\,S_n^k[s_v] - (k+1)\,S_n^{k+1}[s_v] = S_n^{k+1}[v\,a_v]$$

auf den ersten Teil zurückgeführt werden kann. Für sich allein genommen besagt (24) weniger als (23).

Satz 9. *Unter der Umkehrbedingung*

$$s_n = O(1) \quad \text{oder} \quad \int_{-\pi}^{\pi} |\Re F(z)|\,d\varphi = O(1) \tag{25}$$

gilt (A) $\Rightarrow$ (C, k) *für* $k>0$, *aber nicht für* $k=0$.

Satz 10. *Unter der Umkehrbedingung*

$$S_n^\alpha[s_v] = O(n^\alpha) \quad \text{oder} \quad \int_{-\pi}^{\pi} \left|\frac{F(z)}{(1-z)^\alpha}\right| d\varphi = O(1)/(1-r)^\alpha \tag{26}$$

mit $\alpha>0$ *gilt* (A) $\Rightarrow$ (C, k) *für* $k>\alpha$, *aber nicht für* $k=\alpha$.

Satz 11. *Unter der Umkehrbedingung*

$$n\,a_n = O(1) \quad \text{oder} \quad \int_{-\pi}^{\pi} |\Re\, z f'(z)|\,d\varphi = O(1) \tag{27}$$

gilt (A) $\Rightarrow$ (C, k) *für* $k>-1$.

Satz 12. *Unter der Umkehrbedingung*

$$S_n^\alpha[v\,a_v] = O(n^\alpha) \quad \text{oder} \quad \int_{-\pi}^{\pi} \left|\frac{z f'(z)}{(1-z)^\alpha}\right| d\varphi = O(1)/(1-r)^\alpha \tag{28}$$

mit $\alpha>0$ *gilt* (A) $\Rightarrow$ (C, k) *für* $k>\alpha-1$, *aber nicht für* $k=\alpha-1$.

Beweis. Die Sätze 9 und 11 stellen die Grenzfälle der Sätze 10 und 12 für $\alpha=0$ dar. Auf Grund der Cauchy-Fourierschen Koeffizientenformel

$$s_n r^n = \frac{1}{\pi}\int_{-\pi}^{\pi} \Re\,(F(z))\,e^{-in\varphi}\,d\varphi, \qquad n\geq 1,\ r=1-\frac{1}{n}$$

haben diesmal die funktionentheoretischen Umkehrbedingungen die entsprechenden arithmetischen zur Folge. Daher folgt der positive Teil der Sätze 9 bis 12 aus den Sätzen 5 bis 8 in Verbindung mit dem Konvexitätssatz. Die Gegenbeispiele berechnen sich einschließlich des Grenzfalles $\alpha=0$ aus

$$(1-z)^{-\alpha}F(z) = (1+z)^{-\alpha-1} \quad \text{bzw.} \quad (1-z)^{-\alpha}z f'(z) = z(1+z)^{-\alpha-1}.$$

Man beachte, daß $\Re(1+z)^{-1}>0$ ist für $|z|<1$ und daß beim zweiten Beispiel noch a_0 passend zu wählen ist.

Es ist bemerkenswert, daß die Umkehrbedingung der Sätze 6 und 8 die Umkehrbedingungen der Sätze 10 und 12 mit $\alpha+1$ oder $\alpha+1+\varepsilon$ statt α zur Folge haben. Damit wird der positive Teil der Sätze 6 und 8 aus den Sätzen 10 und 12 ableitbar. Die Sätze 10 und 12 können im Fall der funktionentheoretischen Umkehrbedingungen natürlich auch mit unserer funktionentheoretischen Methode direkt hergeleitet werden. Der Beweis wird dabei sogar besonders einfach[3]).

7. Nach dem Studium des zentralen Falles $\varrho=2$ und der Extremfälle $\varrho=1$, $\varrho=+\infty$ wenden wir uns nun der Untersuchung des allgemeinen Falles $1<\varrho<+\infty$ zu. Je nachdem ob $\varrho\leqq 2$ oder $\geqq 2$ ist, treten die funktionentheoretischen oder die arithmetischen Umkehrbedingungen in den Vordergrund. Wir setzen $\frac{1}{\varrho}+\frac{1}{\varrho'}=1$.

Satz 13. *Unter der Umkehrbedingung*

$$(29)\qquad \sum_{\nu=0}^{n}|s_\nu|^\varrho=O(n)\qquad \textit{oder}\qquad \int_{-\pi}^{\pi}|F(z)|^{\varrho'}\,d\varphi=O(1)/(1-r)^{\varrho'-1}$$

gilt $(A)\Rightarrow(C,k)$ *für* $k>\frac{1}{\varrho}$, *aber nicht für* $k=\frac{1}{\varrho}$.

Satz 14. *Unter der Umkehrbedingung*

$$(30)\qquad \sum_{\nu=0}^{n}|S_\nu^\alpha[s_\mu]|^\varrho=O(n^{\varrho\alpha+1})\qquad \textit{oder}\qquad \int_{-\pi}^{\pi}\left|\frac{F(z)}{(1-z)^\alpha}\right|^{\varrho'}d\varphi=O(1)/(1-r)^{\varrho'\alpha+\varrho'-1}$$

mit $\alpha>-\frac{1}{\varrho}$ *gilt* $(A)\Rightarrow(C,k)$ *für* $k>\alpha+\frac{1}{\varrho}$, *aber nicht für* $k=\alpha+\frac{1}{\varrho}$.

Satz 15. *Unter der Umkehrbedingung*

$$(31)\qquad \sum_{\nu=0}^{n}|\nu a_\nu|^\varrho=O(n)\qquad \textit{oder}\qquad \int_{-\pi}^{\pi}|zf'(z)|^{\varrho'}\,d\varphi=O(1)/(1-r)^{\varrho'-1}$$

gilt $(A)\Rightarrow(C,k)$ *für* $k>\frac{1}{\varrho}-1$, *aber nicht für* $k=\frac{1}{\varrho}-1$.

Satz 16. *Unter der Umkehrbedingung*

$$(32)\qquad \sum_{\nu=0}^{n}|S_\nu^\alpha[\mu a_\mu]|^\varrho=O(n^{\varrho\alpha+1})\qquad \textit{oder}\qquad \int_{-\pi}^{\pi}\left|\frac{zf'(z)}{(1-z)^\alpha}\right|^{\varrho'}d\varphi=O(1)/(1-r)^{\varrho'\alpha+\varrho'-1}$$

mit $\alpha>-\frac{1}{\varrho}$ *gilt* $(A)\Rightarrow(C,k)$ *für* $k>\alpha+\frac{1}{\varrho}-1$, *aber nicht für* $k=\alpha+\frac{1}{\varrho}-1$.

[3]) Da der funktionentheoretische Teil des folgenden § 7 auch hierauf zurückgeführt werden kann, so kann hierin wohl der Hauptfall der funktionentheoretischen Methode gesehen werden.

Bemerkenswert sind hierbei die sich aus Satz 14 und 16 ergebenden Grenzfälle:

Unter der Umkehrbedingung

$$(33)\qquad \sum_{\nu=0}^{n} |S_\nu^{-\frac{1}{\varrho}}[s_\mu]|^\varrho = O(1) \qquad \textit{oder} \qquad \int_{-\pi}^{\pi} \frac{|(1-z)F(z)|^{\varrho'}}{|1-z|}\,d\varphi = O(1)$$

gilt $(A) \Rightarrow (C, k)$ *für* $k > 0$, *und unter der Umkehrbedingung*

$$(34)\qquad \sum_{\nu=0}^{n} |S_\nu^{-\frac{1}{\varrho}}[\mu a_\mu]|^\varrho = O(1) \qquad \textit{oder} \qquad \int_{-\pi}^{\pi} \frac{|(1-z)\,z f'(z)|^{\varrho'}}{|1-z|}\,d\varphi = O(1)$$

gilt $(A) \Rightarrow (C, k)$ *für* $k > -1$.

Beweis. Für $\alpha = 0$ geht Satz 14 und 16 in Satz 13 und 15 über. Nach der Ungleichung von HAUSDORFF-YOUNG folgen für $\varrho \leq 2$ die funktionentheoretischen Umkehrbedingungen aus den entsprechenden arithmetischen, während für $\varrho \geq 2$ das Umgekehrte der Fall ist. Nach der Ungleichung von HÖLDER gilt für $k > \alpha + \frac{1}{\varrho} \geq 0$

$$S_n^k[s_\nu] = O(1)\sum_{\nu=0}^{n}(n+1-\nu)^{k-\alpha-1}\,|S_\nu^\alpha[s_\mu]|$$

$$= O\left(n^{k-\alpha-\frac{1}{\varrho}}\right)\left(\sum_{\nu=0}^{n}|S_\nu^\alpha|^\varrho\right)^{1/\varrho},$$

$$\int_{-\pi}^{\pi}\left|\frac{F(z)}{(1-z)^k}\right|d\varphi = \frac{O(1)}{(1-r)^{k-\alpha-\frac{1}{\varrho}}}\left(\int_{-\pi}^{\pi}\left|\frac{F(z)}{(1-z)^\alpha}\right|^{\varrho'}d\varphi\right)^{1/\varrho'}.$$

Daher folgt der positive Teil der Sätze 13 bis 16 einschließlich der Grenzfälle $\alpha = -\frac{1}{\varrho}$ aus den Sätzen 10 und 12. Das Gegenbeispiel für Satz 14 berechnet sich aus

$$S_n^\alpha[s_\nu] = \begin{cases} n^{\alpha+\frac{1}{\varrho}} & \text{für} \quad n = 2^m \quad (m = 0, 1, \ldots) \\ 0 & \text{sonst.} \end{cases}$$

Man beachte, daß in diesem Fall $\sum_{\nu=0}^{n}|S_\nu^\alpha| = O\left(n^{\alpha-\frac{1}{\varrho}}\right)$ und folglich

$$(1-z)^{-\alpha}F(z) = O(1)\,(1-r)^{-\alpha-\frac{1}{\varrho}}$$

ist. Entsprechend berechnet sich das Gegenbeispiel für Satz 16 aus

$$S_n^\alpha[\nu a_\nu] = \begin{cases} n^{\alpha+\frac{1}{\varrho}} & \text{für} \quad n = 2^m \quad (m = 0, 1, \ldots) \\ 0 & \text{sonst,} \end{cases}$$

wobei noch a_0 passend zu bestimmen ist.

8. Wir schließen noch einige Bemerkungen an.

Die bewiesenen Sätze gewinnen nach Ersetzung von $S_n^\alpha[s_\nu]$ durch s_n oft eine übersichtlichere Form. So entsteht z.B. aus Satz 10 der

Satz 17. *Unter der Umkehrbedingung*

$$s_n = O(n^\alpha) \quad \text{oder} \quad \int_{-\pi}^{\pi} |F(z)|\, d\varphi = O(1)/(1-r)^\alpha \tag{35}$$

mit $\alpha > 0$ *folgt aus*

$$(1-z)\,F(z) = o(1)\,(1-z)^{-\alpha} \quad \text{für reelle } z \to 1-0 \tag{36}$$

sogar

$$S_n^k[s_\nu] \Big/ \binom{n+k}{n} = o(n^\alpha) \quad (n \to \infty) \tag{37}$$

für jedes $k > 0$, *aber nicht mehr für* $k = 0$.

In der Sprache der Limitierungstheorie besagen (36) bzw. (37), daß $s_n = o(n^\alpha)$ im Sinne (A) bzw. (c, k) ist. Die Umkehrbedingung (35) ist so gebildet, daß sie für $s_n = \binom{n+\alpha}{n}$ gerade erfüllt ist.

Die funktionentheoretischen Umkehrbedingungen lassen sich nach der Theorie der Fourier-Reihen mit Lebesgueschen Klassen und Lipschitz-Klassen ausdrücken und so mit Eigenschaften der Funktionen entlang des Einheitskreises in Zusammenhang bringen. Gemeinsam mit Herrn A. PEYERIMHOFF ist eine Arbeit in Vorbereitung, die ausgehend von den Funktionenklassen bei Fourier-Reihen die allgemeinen Zusammenhänge zwischen den funktionentheoretischen und den arithmetischen Bedingungen klärt und die Bedingungen selbst in ihrer gegenseitigen Abhängigkeit voneinander untersucht. Hierhin gehört z.B. die Frage, wie sich die Umkehrbedingung (30) für verschiedene α, ϱ mit gleichem $\alpha + \frac{1}{\varrho}$ (d.h. gleicher Umkehrwirkung) verhält. Es läßt sich zeigen, daß dabei die arithmetische Umkehrbedingung für größer werdendes ϱ schwächer wird.

Die Verwendung der Parsevalschen Gleichung erlaubt auch die Verbesserung einer Abschätzung bei einem gemeinsam mit den Herren R. BOJANIĆ und A. PEYERIMHOFF behandelten, verwandten Problem [1]:

Satz 18. *Aus* $a_n \geq 0$ *und* $\sum_{\nu=0}^{n} a_\nu s_{n-\nu} = \frac{n^2}{2} + O(n)$ *folgt* $s_n = n + O(n^{\frac{3}{4}})$.

Das Restglied enthielt vorher noch den Faktor $\log^{\frac{1}{4}} n$ [4]). Beim Beweis setzt man $(1-z)F^2(z) = (1-z)^{-3} + G(z)$ und benutzt jetzt neben $G(z) = O(1)\,(1-|z|)^{-2}$ die schärfere Abschätzung $\int_{-\pi}^{\pi} |G(z)|^2 d\varphi = O(1)\,(1-r)^{-3}$. Für genügend kleines, aber festes $\delta > 0$ gilt somit

$$(1-z)^2 F(z) - 1 = \begin{cases} O(1)\,|1-z|^3 \cdot |G(z)| & \text{für } |1-z| \leq \delta(1-r)^{\frac{3}{4}} \\ O(1)\,|1-z|^{\frac{3}{4}}(1-r)^{-1} & \text{für } |1-z| \geq \delta(1-r)^{\frac{3}{4}} \end{cases}$$

[4]) Herr H. E. RICHERT teilte mir freundlicherweise mit, daß er unabhängig die gleiche Verbesserung gefunden hat.

und daher für $r = 1 - \frac{1}{n}$ $(n \to \infty)$ nach CAUCHY-SCHWARTZ[5])

$$S_n^2[a_\nu - 1] = O(1) \int\limits_{|\varphi| \le \delta n^{-\frac{1}{2}}} |G(z)| \cdot |1 - z|^{-1} d\varphi + O(1) \int\limits_{\delta n^{-\frac{1}{2}} \le |\varphi| \le \pi} n |\varphi|^{-\frac{1}{2}} d\varphi$$

$$= O(1)\, n^{\frac{1}{2}} \cdot n^{\frac{1}{2}} + O(1)\, n \cdot n^{\frac{1}{2} \cdot \frac{1}{2}} = O(n^2).$$

Wegen $a_\nu - 1 \geq -1$ folgt aus Konvexitätsgründen $S_n^1[a_\nu - 1] = O(n^{\frac{1}{2}})$.

Die in dieser Arbeit angegebenen Beweismethoden sind nicht auf Potenzreihen beschränkt, sondern lassen sich allgemeiner auch auf Dirichlet-Reihen und Laplace-Transformationen sinngemäß ausdehnen. Die bewiesenen Sätze sind mit der arithmetischen Umkehrbedingung zu einem Teil schon bekannt gewesen: Der positive Teil von Satz 15 (Satz 2 und 7) findet sich im wesentlichen schon in [4, Lemma 4], mit der diese Arbeit auch sonst manche Berührungspunkte hat. Für eine Verallgemeinerung dieses Satzes auf Dirichlet-Reihen vgl. [9]. Die arithmetische Umkehrbedingung (31) verbessert die bekannte Umkehrbedingung $\sum n |a_n|^2 < \infty$ von FEJÉR [vgl. 8, S. 65] und ihre Verallgemeinerungen $\sum n^{\varrho-1} |a_n|^\varrho < \infty$ in [3]. Außerdem ist der arithmetische Teil von § 6 als bekannt anzusehen.

Literatur

[1] BOJANIĆ, R., W. JURKAT u. A. PEYERIMHOFF: Über einen Tauber-Satz für Faltungen. Math. Z. **65**, 195—199 (1956). — [2] HARDY, G. H.: Divergent series. Oxford 1949. — [3] HARDY, G. H., and J. E. LITTLEWOOD: Some theorems concerning DIRICHLET's series. Messenger of Math. **43**, 134—147 (1914). — [4] HARDY, G. H., u. J. E. LITTLEWOOD: A convergence criterion for Fourier series. Math. Z. **28**, 612—634 (1928). — [5] IZUMI, S.: A simple proof of LITTLEWOOD's Tauberian theorem. Proc. Japan Acad. **30**, 927—929 (1954). — [6] JURKAT, W. B.: Ein funktionentheoretischer Beweis für O-Tauber-Sätze bei Potenzreihen. Archiv d. Math. **7**, 122—125 (1956). — [7] JURKAT, W. B.: Ein funktionentheoretischer Beweis für O-Tauber-Sätze bei den Verfahren von BOREL und EULER-KNOPP. Archiv d. Math. **7**, 278—283 (1956). — [8] LANDAU, E.: Darstellung und Begründung einiger neuerer Ergebnisse der Funktionentheorie, 2. Aufl. Berlin 1929. — [9] SZÁSZ, O.: Verallgemeinerung und neuer Beweis einiger Sätze Tauberscher Art. Sitzgsber. München **59**, 325—340 (1929). — [10] TITCHMARSH, E. C.: The theory of functions. Oxford 1932. — [11] WIELANDT, H.: Zur Umkehrung des Abelschen Stetigkeitssatzes. Math. Z. **56**, 206—207 (1952).

[5]) Hiermit ist die einzige Veränderung gegenüber dem früheren Beweis verbunden.

Dept. of Mathematics, Ohio State University, Columbus 10, Ohio (USA)

(Eingegangen am 4. September 1956)

GRUNSKY, H.
Math. Zeitschr. Bd. 67, S. 223—228 (1957)

Über konforme Abbildungen, die gewisse Gebietsfunktionen in elementare Funktionen transformieren. II

Von
HELMUT GRUNSKY

Die in Teil I[1]) verwendete Methode läßt sich so weiterbilden, daß sie auch zum Beweise der dort in Anmerkung [2]) erwähnten anderen Ergebnisse von DE LA VALLÉE POUSSIN und JULIA ausreicht. Es dürfte zweckmäßig sein, das Beweisprinzip zunächst in Gestalt eines allgemeinen Hilfssatzes zu formulieren.

Unter einer linken oder rechten Halbumgebung eines analytischen Jordan-Bogens Z in der Ebene werde ein beliebiges, teilweise von Z berandetes Gebiet verstanden, das zu jedem Punkt von Z eine Halbumgebung enthält und das bei einer Durchlaufung des Bogens im Sinne seiner Orientierung zur Linken bzw. Rechten liegt.

Hilfssatz. Es seien Z_1 und Z_2 zwei abgeschlossene orientierte analytische Jordan-Bogen in der Ebene der komplexen Veränderlichen z, $\mathfrak{Z}_1$ eine linke, $\mathfrak{Z}_2$ eine rechte Halbumgebung von Z_1 bzw. Z_2. $g_\iota(z)$, $\iota = 1, 2$, sei harmonisch und nicht konstant in $\mathfrak{Z}_\iota \cup Z_\iota$ und auf Z_ι gelte, wenn $\partial/\partial n$ die Ableitung nach der rechts weisenden Normalen bedeutet:

$$g_\iota(z) = c = \text{const}, \quad \frac{\partial g_\iota}{\partial n} \geq 0.$$

Ohne Beschränkung der Allgemeinheit werde ferner angenommen

$$\int_{Z_1} \frac{\partial g_1}{\partial n}\, ds = \int_{Z_2} \frac{\partial g_2}{\partial n}\, ds \qquad (s\text{: Bogenlänge auf } Z_\iota)$$

(die Gültigkeit dieser Voraussetzung läßt sich durch passende Verkürzung des einen Bogens erzwingen). Dann gibt es eine punktweise Identifizierung der Bogen Z_1 und Z_2 miteinander, derart, daß die auf der so entstandenen Riemannschen Mannigfaltigkeit $\mathfrak{Z}$ durch $g_\iota(z)$ $(\iota = 1, 2)$ erklärte Funktion überall auf $\mathfrak{Z}$ harmonisch ist.

Beweis. Der Anfangspunkt von Z_ι sei $z_0^{(\iota)}$. Durch

$$\varphi_\iota(z^{(\iota)}) = \int_{z_0^{(\iota)}}^{z^{(\iota)}} \frac{\partial g_\iota}{\partial n}\, ds, \qquad z^{(\iota)} \in Z_\iota$$

wird auf Z_ι eine im strengen Sinne monoton wachsende Funktion der Bogenlänge erklärt $\left(\frac{\partial g_\iota}{\partial n} = 0 \text{ kann nur in endlich vielen Punkten auf } Z_\iota \text{ gelten}\right)$.

[1]) Teil I s. Math. Z. 67, 129—132 (1957).

Wir identifizieren zwei Punkte $z^{(i)}$, wenn $\varphi_1(z^{(1)}) = \varphi_2(z^{(2)})$. Die entstandene Riemannsche Mannigfaltigkeit bilden wir umkehrbar eindeutig und konform auf ein schlichtes Gebiet $\mathfrak{W}$ der w-Ebene ab; $\mathfrak{W}$ sei das Bild von $\mathfrak{Z}_i$, W das gemeinsame Bild der Z_i, w_0 das der $z_0^{(i)}$. Durch Verpflanzung der $\mathfrak{g}_i(z)$ entsteht eine stetige Funktion $\tilde{\mathfrak{g}}(w)$, die zunächst sicher in allen nicht auf W liegenden Punkten harmonisch ist. Die Verpflanzung der $\varphi_i(z^{(i)})$ ergebe $\tilde{\varphi}_i(w)$ auf W. Wegen der Invarianz von $\frac{\partial \mathfrak{g}_i}{\partial n}\,ds$ gegenüber konformer Abbildung ist

$$\tilde{\varphi}_i(w) = \int_{w_0}^{w} \frac{\partial \tilde{\mathfrak{g}}_i}{\partial n}\,ds \quad (s\text{: Bogenlänge auf } W).$$

Gemäß der Identifizierungsvorschrift ist aber $\tilde{\varphi}_1(w) = \tilde{\varphi}_2(w)$, also auch $\frac{\partial \tilde{\varphi}_1}{\partial s} = \frac{\partial \tilde{\varphi}_2}{\partial s}$, d.h. $\frac{\partial \tilde{\mathfrak{g}}_1}{\partial n} = \frac{\partial \tilde{\mathfrak{g}}_2}{\partial n}$. Die Funktion $\tilde{\mathfrak{g}}(w)$ ist daher auf W nicht nur stetig, sondern auch mit stetigen Ableitungen versehen, und daraus folgt, daß sie auch dort noch harmonisch ist[2]).

Es sei nun $\mathfrak{Z}$ ein Gebiet in der z-Ebene der endlichen Zusammenhangszahl $n > 1$ und $\infty \in \mathfrak{Z}$. Sein Rand Z bestehe aus einfach geschlossenen analytischen Kurven Z_ν $(\nu = 1, 2, \ldots, n)$. Ferner setze man wie in Teil I:

$$\mathfrak{g}(z) = \sum_{\varkappa=1}^{k} \alpha_\varkappa g(z, z_\varkappa) + \sum_{\nu=1}^{n} c_\nu h_\nu(z), \tag{1}$$

wo $g(z, z_\varkappa)$ die Greensche Funktion von $\mathfrak{Z}$ mit Singularität $z_\varkappa \neq \infty$, $h_\nu(z)$ das harmonische Maß von Z_ν bedeutet; die Punkte $z_\varkappa \in \mathfrak{Z}$ und die reellen Konstanten $\alpha_\varkappa$ und c_ν sollen jetzt keinerlei Einschränkungen unterworfen sein, d.h. die Bedingung 3. aus Teil I fällt weg. $\mathfrak{g}(z)$ besitzt auf jeder Randkomponenten Z_ν eine gerade Anzahl von kritischen Stellen mit Vorzeichenwechsel der nach der äußeren Normalen genommenen Ableitung $\frac{\partial \mathfrak{g}}{\partial n}$. Es seien Z' und Z'' zwei im Punkt z_0 zusammenstoßende Bogen einer Randkomponenten Z_ν, auf denen $\frac{\partial \mathfrak{g}}{\partial n}$ verschiedenes Vorzeichen hat. Orientieren wir Z' und Z'' indem wir für beide z_0 als Anfangspunkt festsetzen, und verstehen wir unter $\frac{\partial}{\partial n}$ wieder die Ableitung in Richtung der von dem orientierten Bogen nach rechts weisenden Normalen, so können wir unseren Hilfssatz anwenden, wobei z_0 die Rolle von $z_0^{(1)}$ und $z_0^{(2)}$ übernimmt, je ein (nicht notwendig echter) Teilbogen von Z' und Z'' die von Z_1 bzw. Z_2. Man kann die Identifizierung

[2]) Das folgt leicht aus dem Satz: Eine stetig differenzierbare Funktion $f(w)$ ist harmonisch, wenn $\int_K \frac{\partial f}{\partial n}\,ds = 0$ über den Rand K jedes im Definitionsgebiet liegenden Kreises (vgl. KELLOGG, Foundations of Potential Theory, Berlin 1929, Kap. 8, Satz XIII, S. 227). Diese Bedingung ist in unserem Falle erfüllt, auch für jeden Kreis $\mathfrak{K}$, der teilweise zu $\mathfrak{W}_1$, teilweise zu $\mathfrak{W}_2$ gehört, denn das Integral verschwindet, wenn es über den Rand von $\mathfrak{K} \cap \mathfrak{W}_i$ erstreckt wird, und die Summe dieser Integrale ist das Integral über K, da sich die über Bogen von W erstreckten Teilintegrale herausheben.

so weit treiben, daß einer der beiden Bogen Z' und Z'' vollständig erfaßt wird. Bildet man die so erklärte Riemannsche Mannigfaltigkeit schlicht in die w-Ebene ab, so erhält man ein Gebiet $\mathfrak{W}^*$ der Zusammenhangszahl n oder $n-1$; der zweite Fall tritt ein, wenn $Z' \cup Z'' = Z_\nu$ und wenn $\int_{Z_\nu} \frac{\partial \mathfrak{g}}{\partial n}\, ds = 0$. Das Bild $\mathfrak{W}$ von $\mathfrak{Z}$ selbst entsteht aus $\mathfrak{W}^*$, indem man einen Schlitz oder Einschnitt $W_\nu^{(0)}$ längs des Bildes der beiden identifizierten Teilbogen von Z' und Z'' führt. Die durch Verpflanzung von $\mathfrak{g}(z)$ entstandene Funktion $\tilde{\mathfrak{g}}(w)$ hat in bezug auf $\mathfrak{W}^*$ dieselben Eigenschaften wie $\mathfrak{g}(z)$ in bezug auf $\mathfrak{Z}$, jedoch ist die Zahl der Vorzeichenwechsel von $\frac{\partial \tilde{\mathfrak{g}}}{\partial n}$ auf dem Rande von $\mathfrak{W}^*$ um zwei geringer; außerdem gilt $\tilde{\mathfrak{g}}(w) = c_\nu$ auch auf $W_\nu^{(0)}$ und es ist für alle ν:

$$\int_{W_\nu^*} \frac{\partial \tilde{\mathfrak{g}}}{\partial n}\, ds = \int_{Z_\nu} \frac{\partial \mathfrak{g}}{\partial n}\, ds,$$

wo W_ν^* diejenige Randkomponente von $\mathfrak{W}^*$ bedeutet, die dem Bilde von Z_ν angehört.

Auf $\mathfrak{W}^*$ und $\tilde{\mathfrak{g}}$ wenden wir nötigenfalls dieselben Überlegungen an, wie soeben auf $\mathfrak{Z}$ und $\mathfrak{g}$ und erhalten ein Gebiet $\mathfrak{W}^{**}$ und eine Funktion $\tilde{\tilde{\mathfrak{g}}}$. Dabei kann es sein, daß der Endpunkt $w_\nu^{(0)}$ des Schlitzes $W_\nu^{(0)}$ auf dem Rande von $\mathfrak{W}^*$ innerer Punkt eines zu verheftenden Bogens ist. In diesem Falle geht das neue Bildgebiet von $\mathfrak{Z}$ aus $\mathfrak{W}^{**}$ hervor, indem man einen verzweigten Einschnitt führt[3]).

Nach endlich vielen Schritten kommt man zu einem Gebiet und einer Funktion — sie seien wieder mit $\mathfrak{W}^*$ und $\tilde{\mathfrak{g}}$ bezeichnet —, für die folgendes gilt: Die Anzahl m der Randkomponenten von $\mathfrak{W}^*$ ist $\leq n$. Das Bild $\mathfrak{W}$ von $\mathfrak{Z}$ entsteht aus $\mathfrak{W}^*$, indem man gewisse, im allgemeinen verzweigte, Einschnitte und Schlitze anbringt, deren jeder auf einer Niveaulinie von $\tilde{\mathfrak{g}}$ liegt. Diese Funktion ist in $\mathfrak{W}^*$ überall harmonisch bis auf die Bilder $w_\varkappa$ der Punkte $z_\varkappa$, in denen sie sich verhält wie $-\alpha_\varkappa \log|w-w_\varkappa| + O(1)$. $\frac{\partial \tilde{\mathfrak{g}}}{\partial n}$ hat auf keiner der eventuellen Randkomponenten W_ν^* von $\mathfrak{W}^*$ einen Vorzeichenwechsel; es ist $\tilde{\mathfrak{g}}(w) = c_\nu$ auf W_ν^*, $\int_{W_\nu^*} \frac{\partial \tilde{\mathfrak{g}}}{\partial n}\, ds = \int_{Z_\nu} \frac{\partial \mathfrak{g}}{\partial n}\, ds$. Im Falle $m = 0$ ist (bei passender Normierung) $\tilde{\mathfrak{g}}(w)$ schon eine elementare Funktion von der im nachstehenden Satze gekennzeichneten Art; im Falle $m > 0$ können wir Satz 1 aus Teil I anwenden.

Der geschilderte Prozeß wird im allgemeinen zu einem anderen Endergebnis führen, wenn er mit einem andern Trennpunkt zwischen zwei Teilbogen verschiedenen Vorzeichens von $\frac{\partial \mathfrak{g}}{\partial n}$ begonnen wird, oder wenn er, falls möglich, bei einem späteren Schritt in entsprechender Weise abgewandelt

[3]) Eine zusammenhängende Vereinigungsmenge von endlich vielen Jordan-Bogen, die keinen geschlossenen Weg enthält, heiße verzweigt, wenn sie nicht wieder selbst einen Jordan-Bogen darstellt.

wird; jedoch sind die Bedingungen für die Möglichkeit solcher Abänderungen kaum zu übersehen. Indessen gibt es noch eine andere Art, den Prozeß zu variieren, die bisher nicht in Betracht gezogen wurde. Es ist zwar immer möglich, aber nicht notwendig, bei einem Schritt den einen der beiden Bogen Z' und Z'' vollständig aufzubrauchen. Man kann mit der Identifizierung aufhören, wenn von beiden Bogen, ausgehend von z_0, nur je ein echter Teilbogen erfaßt wurde. Dann liegt allerdings keine Reduktion vor, die Anzahl der Vorzeichenwechsel von $\frac{\partial \mathfrak{g}}{\partial n}$ bleibt dieselbe. Indes kann man, von dem erreichten Zwischenstadium ausgehend, in endlich vielen Schritten zu Ende kommen und wird dabei im allgemeinen ein anderes Ergebnis erhalten als vorhin, von gewissen Ausnahmefällen abgesehen; nämlich, wenn im weiteren Verlauf unvermeidlich einmal das Bild $w_\nu^{(0)}$ der beiden Punkte z' und z'', in denen der erste Schritt abgebrochen wurde, und die von $w_\nu^{(0)}$ ausgehenden Bilder der Reste der Bogen Z' und Z'' in der Rolle von z_0, Z', Z'' des geschilderten ersten Schrittes auftreten; denn dann war ja das Abbrechen des ersten Schrittes nur scheinbar. Diese Ausnahme liegt aber genau in dem schon einmal erwähnten Fall vor, bei dem auf Z_ν zwei Zeichenwechsel stattfinden und außerdem $\int_{Z_\nu} \frac{\partial \mathfrak{g}}{\partial n} ds = 0$ ist. In jedem anderen Fall können z' und z'' so gewählt werden, daß einer der beiden Restbogen von Z' und Z'', sagen wir der von Z', mit seinem vom Rest von Z'' verschiedenen Nachbarbogen verheftet werden kann, derart, daß jetzt $w_\nu^{(0)}$ noch mit einem weiteren Randpunkt identifiziert und damit zu einem Verzweigungspunkt eines Einschnittes oder Schlitzes wird.

Wir wollen noch vereinbaren, daß jede der benutzten Abbildungen den Punkt ∞ festlasse und dort den Streckungsfaktor 1 und Drehung und Verschiebung 0 habe. Dann ist bis auf die Einzigkeitsaussage folgender Satz bewiesen:

Satz 2. *Sei $\mathfrak{Z}$ ein in der z-Ebene liegendes Gebiet der endlichen Zusammenhangszahl $n > 1$, dessen Randkomponenten Z_ν alle echte Kontinuen seien und das ∞ enthalte; es gibt eine konforme Abbildung $w = F(z)$, die sich bei ∞ verhält wie $z + O\left(\frac{1}{z}\right)$, bei der $\mathfrak{g}(z)$ [vgl. (1)] transformiert wird in*

$$\tilde{\mathfrak{g}}(w) = \log |S(w)|$$

mit

$$S(w) = A \prod_{\varkappa=1}^{k} (w - w_\varkappa)^{-\alpha_\varkappa} \prod_{\nu=1}^{n} (w - b_\nu)^{-\beta_\nu},$$

$$w_\varkappa = F(z_\varkappa), \quad \beta_\nu = \int_{Z_\nu} \frac{\partial \mathfrak{g}}{\partial n} ds, \quad A = e^{\mathfrak{g}(\infty)}.$$

In jeder Komponenten der Komplementärmenge des Bildgebietes mit $\beta_\nu \neq 0$ liegt genau ein Punkt b_ν. Die Abbildung ist dann und nur dann einzig, wenn $\frac{\partial \mathfrak{g}}{\partial n}$ auf jeder Randkomponenten höchstens zwei Vorzeichenwechsel aufweist und $\int_{Z_\nu} \frac{\partial \mathfrak{g}}{\partial n} ds = 0$ gilt für alle Z_ν, auf denen Zeichenwechsel vorkommen.

Der Beweis der Einzigkeitsaussage ergibt sich wie bei Satz 1. Ist auf irgendeine Weise eine Abbildung von der in dem Satz geschilderten Art gefunden, so lege man das Bildgebiet $\mathfrak{Z}'$ in die z'-Ebene und wende auf dieses den geschilderten Prozeß an. Wegen der Invarianz von $\frac{\partial g}{\partial n}ds$ ist das Ergebnis dasselbe wie wenn wir von $\mathfrak{Z}$ ausgehen. In dem unserer Einzigkeitsaussage zugrunde liegenden Fall ist der Verlauf ein wohlbestimmter und es werden immer nur Punkte identifiziert, die schon in der z'-Ebene zusammenfallen. Daher können alle auftretenden Abbildungsfunktionen jeweils in der ganzen Ebene regulär gewählt werden und sie sind dann mit Rücksicht auf die Normierung Identitäten. Das heißt aber, daß jeder Bildbereich von der in unserem Satz genannten Art identisch ist mit dem durch unseren Prozeß gewonnenen.

Wir wählen nun speziell

$$g(z) = \sum_{\nu=1}^{n-1} c_\nu h_\nu(z) \tag{2}$$

und bestimmen die c_ν so, daß

$$\beta_\nu = \frac{1}{2\pi}\int_{Z_\nu} \frac{\partial g}{\partial n}\,ds = -1, \quad \nu = 1, \ldots, n-1. \tag{3}$$

Diese Forderungen führen bekanntlich auf ein System linearer Gleichungen mit nicht verschwindender Determinanten für die c_ν, $\nu = 1, \ldots, n-1$ und lassen sich also auf genau eine Weise erfüllen. Wegen $\sum_{\nu=1}^{n} \beta_\nu = 0$ ergibt sich ferner:

$$\beta_n = \frac{1}{2\pi}\int_{Z_n} \frac{\partial g}{\partial n}\,ds = n-1. \tag{4}$$

Dabei sind die $c_\nu < 0$, denn wäre etwa $c_1 \geq 0$ das größte unter den c_ν, so wäre $g(z)$ in $\mathfrak{Z}$ überall kleiner als auf Z_1, und hier also $\frac{\partial g}{\partial n} \geq 0$, im Widerspruch zu (3). Daher ist auf Z_n überall $\frac{\partial g}{\partial n} > 0$. Normieren wir die Abbildung in die w-Ebene diesmal so, daß $b_n = \infty$, so ist das Bild W_n von Z_n eine einfach geschlossene Kurve, $S(w)$ ein Polynom $(n-1)$-ten Grades mit den Nullstellen $b_1, \ldots, b_{n-1}$, und $|S(w)|$ ist im Äußern von W_n größer, im Innern kleiner als auf W_n, d.h. diese Kurve ist eine vollständige Niveaulinie von $S(w)$. Ebenso sieht man, daß eine Randkomponente, für die c_ν minimal ist, ein einfach geschlossenes Bild hat. Damit ist bewiesen:

Satz 3 (von JULIA[4]). *Jedes schlichte Gebiet der endlichen Zusammenhangszahl $n > 1$ ohne einpunktige Randkomponente läßt sich umkehrbar eindeutig und konform auf ein beschränktes Gebiet abbilden, dessen Randkomponenten W_ν, $\nu = 1, \ldots, n$ je auf einer Niveaulinie eines von ν unabhängigen Polynoms vom Grade $n-1$ liegen, wobei die äußere Randkomponente eine vollständige Niveaulinie, mindestens eine innere Randkomponente einfach geschlossen ist.*

[4]) Vgl. JULIA, G.: Sur une représentation conforme canonique des aires multiplement connexes. C. r. Acad. Sci. Paris **194**, 819—822 (1932).

Von weiteren Anwendungsmöglichkeiten des Beweisprinzips greifen wir nur noch eine heraus. Wir bestimmen die Koeffizienten in (2) so, daß $\beta_\nu = \pm 1$ auf je $\left[\frac{n}{2}\right]$ der Randkomponenten und, bei ungeradem n, $\beta_\nu = 0$ auf der einen verbleibenden Komponenten. Man erhält für $S(w)$ in beiden Fällen eine rationale Funktion vom Grade $\left[\frac{n}{2}\right]$. Bei ungeradem n geht die Randkomponente mit $\beta_\nu = 0$ in einen, möglicherweise verzweigten, Schlitz über, und in beiden Fällen sind die Bilder der Randkomponenten mit minimalem und maximalem c_ν einfach geschlossene Kurven. Es gilt also der

Satz 4. *Jedes schlichte Gebiet der endlichen Zusammenhangszahl $n > 1$ ohne einpunktige Randkomponente läßt sich umkehrbar eindeutig und konform auf ein beschränktes Gebiet abbilden, dessen Randkomponenten je auf einer Niveaulinie einer durch das Gebiet bestimmten rationalen Funktion vom Grade $\left[\frac{n}{2}\right]$ liegen. Bei ungeradem n ist genau eine der Randkomponenten des Bildes ein (möglicherweise verzweigter) Schlitz, bei geradem n keine; in beiden Fällen sind mindestens zwei Randkomponenten einfach geschlossene Kurven.*

Mainz, Mathematisches Institut der Universität

(Eingegangen am 16. November 1956)

LUDFORD, G. S. S., and S. H. SCHOT
Math. Zeitschr. Bd. 67, S. 229—237 (1957)

On Sonic Limit Lines in the Hodograph Method*)

By

G. S. S. LUDFORD and S. H. SCHOT

It is well-known [5] that in the theory of plane compressible flow two types of limit lines can be distinguished. In a recent paper in this journal HILDA GEIRINGER [7] has identified a third type of limit line, distinct from the first two. She has called such a line a sonic limit line and has stated its properties, in part without proof.

The only known example of a flow with a sonic limit line is the familiar radial (source) flow found by G. I. TAYLOR [1]. Later RINGLEB [3] stated erroneously that this is the only possible flow exhibiting such a limit line, and the subject is not even mentioned in textbook treatments of limit lines (see e.g. LIGHTHILL [6]).

The object of this paper is (i) to give our original derivations of those unproved results in [7] which were quoted by GEIRINGER from a private communication of the present authors, and (ii) to show that there exist other flows containing sonic limit lines. In particular we explicitly construct an example of a flow with a sonic limit line which has no source-like character. This example is a linear combination of two known solutions, one discussed by RINGLEB [2], and the other by TEMPLE and YARWOOD [4]. Like their solutions, the present one is a compressible analogue of the incompressible flow around a semi-infinite plate. In addition to a sonic limit line it contains an example (believed to be the first) of GEIRINGER's "sonic double limit point".

1. Sonic Limit Singularities. We start by giving the basic equations and relations in the theory of the hodograph method, needed in the later discussion.

In the hodograph variables q, ϑ the so-called potential and stream functions φ, ψ satisfy the Chaplygin equations

$$\frac{\partial \varphi}{\partial \vartheta} = \frac{q}{\varrho} \frac{\partial \psi}{\partial q}; \qquad (M^2 - 1) \frac{\partial \psi}{\partial \vartheta} = \varrho q \frac{\partial \varphi}{\partial q}, \tag{1}$$

where

$$\varrho = (1 - q^2)^{\frac{1}{\gamma - 1}}, \tag{2}$$

and

$$M^2 = \frac{2}{\gamma - 1} \cdot \frac{q^2}{1 - q^2} \tag{3}$$

The sonic speed, i.e. the value of q for which $M = 1$, is $q_t = \sqrt{(\gamma - 1)/(\gamma + 1)}$.

*) This work was supported by the Office of Ordnance Research, U.S. Army, under Contract DA—36—034—ORD—1486 with the University of Maryland.

From any non-constant solution $\varphi(q, \vartheta)$, $\psi(q, \vartheta)$ of these equations the coordinates $x(q, \vartheta)$, $y(q, \vartheta)$ are determined by the equations

$$(4) \qquad \begin{cases} dx = \frac{1}{q}\left[\cos\vartheta\, d\varphi - \frac{1}{\varrho}\sin\vartheta\, d\psi\right], \\ dy = \frac{1}{q}\left[\sin\vartheta\, d\varphi + \frac{1}{\varrho}\cos\vartheta\, d\psi\right], \end{cases}$$

the equations (1) being integrability conditions for these.

The hodograph method consists in finding twice continuously differentiable solutions of (1) and determining the corresponding flow in the physical plane defined by (4).

The Jacobian of the transformation from the hodograph plane to the physical plane is

$$(5) \qquad I = \frac{\partial(x, y)}{\partial(q, \vartheta)} = \frac{\partial(x, y)}{\partial(\varphi, \psi)} \cdot \frac{\partial(\varphi, \psi)}{\partial(q, \vartheta)} = -\frac{1}{\varrho^2 q^3}\left[q^2\psi_q^2 + (1 - M^2)\psi_\vartheta^2\right].$$

When $M > 1$ we introduce characteristic coordinates ξ, η of (1) defined by

$$(6) \qquad 2\xi = Q - \vartheta; \quad 2\eta = Q + \vartheta,$$

where

$$Q = \int^q \frac{dq}{q\tan\alpha},$$

and α is the Mach angle $\sin^{-1}\frac{1}{M}$. Then we find

$$(7) \qquad I = -\frac{1}{\varrho^2 q^3}\cot^2\alpha\,\psi_\xi\psi_\eta = \frac{1}{q^3}\varphi_\xi\varphi_\eta.$$

Breakdown of the hodograph transformation will occur at points in the hodograph plane where $I = 0$ and such points are called *limit singularities*. Curves consisting of such points are called *critical curves* and their images in the physical plane are called *limit lines*. It may be shown that limit lines occur only when $M \geq 1$ and that the solution in the neighborhood of these lines has a characteristic geometry.

If $M > 1$ (i.e. $\alpha \neq 90°$) then $I = 0$ implies that either ψ_ξ or ψ_η is zero; the vanishing of both simultaneously can only occur at an isolated point, which is called a *double limit point*. A curve along which $\psi_\xi = 0$, $\psi_\eta \neq 0$ is called a *limit line* L_1; similarly one along which $\psi_\xi \neq 0$, $\psi_\eta = 0$ is called a *limit line* L_2. These terms may still be used when $M = 1$ at some points of the curve.

If $M = 1$ however (i.e. $\alpha = 90°$), I can vanish without either ψ_ξ or ψ_η vanishing. The following classification of sonic limit singularities is due to GEIRINGER

Case		Examples
1) $\psi_\xi \neq 0$ $\psi_\eta \neq 0$	a) $\varphi_\xi = 0$ $\varphi_\eta \neq 0$	
	b) $\varphi_\xi \neq 0$ $\varphi_\eta = 0$	
	c) $\varphi_\xi = 0$ $\varphi_\eta = 0$	radial flow, section 4—sonic limit line
2) $\psi_\xi = 0$ $\psi_\eta \neq 0$	a) $\varphi_\xi = 0$ $\varphi_\eta \neq 0$	
	b) $\varphi_\xi = 0$ $\varphi_\eta = 0$	Ringleb flow — sonic point of an L_1
3) $\psi_\xi \neq 0$ $\psi_\eta = 0$	a) $\varphi_\xi \neq 0$ $\varphi_\eta = 0$	
	b) $\varphi_\xi = 0$ $\varphi_\eta = 0$	Ringleb flow — sonic point of an L_2
4) $\psi_\xi = 0$ $\psi_\eta = 0$	$\varphi_\xi = 0$ $\varphi_\eta = 0$	section 4 — sonic double limit point

(Since the ξ, η-coordinates degenerate on the sonic circle, a derivative with respect to ξ or η for $M = 1$ must be interpreted in the sense of a limit as $M \to 1$ from above.)

In this paper we restrict ourselves to the case $M = 1$, and in particular we prove that only case 1c) can occur along a whole segment of the sonic line S. The image of such a line in the physical plane will be termed a *sonic limit line*, L_t.

2. The Sonic Limit Line. We shall now determine those sonic limit singularities which can occur at isolated points only, and not along a whole non-degenerate arc of the sonic line S.

Theorem. In any region of the q, ϑ-plane where $\psi \in C^2$ and is not constant, the cases 1a), 1b), 2), 3), and 4) of the above classification cannot occur along a segment of the sonic line.

Proof. Consider a point at which one but not both $\varphi_\xi, \varphi_\eta$ are zero. Then

$$\psi_\vartheta = \frac{\varrho}{2} \tan \alpha (\varphi_\xi + \varphi_\eta) \to \infty \quad \text{as} \quad \alpha \to 90°.$$

Since this contradicts the hypothesis $\psi \in C^2$, cases 1a), 1b), 2a), and 3a) cannot occur along a line.

With reference to case 2b) we note that

$$\psi_\xi = q \tan \alpha \, \psi_q - \psi_\vartheta \to 0$$

implies that $\psi_q \to 0$ at the sonic speed $q = q_t = \sqrt{(\gamma - 1)/(\gamma + 1)}$. But then from the TAYLOR formula applied to ψ_q we see that $\psi_q = O(q - q_t)$ while from BERNOULLI's equation in the form (3) we find $\tan \alpha = O(q - q_t)^{-\frac{1}{2}}$. Hence $\psi_q \tan \alpha = O(q - q_t)^{\frac{1}{2}} \to 0$ and $\psi_\vartheta \to 0$. Thus along such a segment $\psi_q = \psi_\vartheta = 0$, which implies $\psi = \text{const}$ in some neighborhood.

A similar argument holds for cases 3b) and 4).

The only remaining type of sonic limit singularity which can occur along a curve is 1c). We shall see that even this case must be broken down into two subcases, of which one can occur only at an isolated point and the other provides the true sonic limit line.

The case 1c) is characterized by $\psi_\xi \neq 0$, $\psi_\eta \neq 0$, $\varphi_\xi = 0$, $\varphi_\eta = 0$. From the last two conditions there follows $\psi_q = 0$, $\psi_\vartheta \cot\alpha = 0$. Hence either $\psi_q = 0$, $\psi_\vartheta = 0$ or $\psi_q = 0$, $\psi_\vartheta \neq 0$ (but such that $\psi_\vartheta \cot\alpha \to 0$ as $\alpha \to 90°$). But when ψ_q, ψ_ϑ are zero along a line, $\psi = \text{const}$ in some neighborhood, and this case is excluded.

These results identify the only type of sonic limit singularity which can provide the true sonic limit line.

Corollary. The necessary and sufficient condition for a curve to be a sonic limit line is

$$M = 1 : \quad \psi_q = 0, \quad \psi_\vartheta \neq 0,$$

at each point.

The sufficiency of this condition follows immediately from (5).

Corollary. Neither an L_1 nor an L_2 can coincide with the image of the sonic line for a whole segment.

3. Geometry of the Sonic Limit Line. The geometry in the neighborhood of the L_t may be discussed just as in the case of the L_1 and L_2 by making use of the concept of "exceptional" directions. An exceptional direction in our case is a direction in the physical plane into which all directions but one at the corresponding point in the hodograph plane map.

From the definition of the L_t we have that

$$d\varphi = \varphi_q\, dq + \varphi_\vartheta\, d\vartheta = 0,$$
$$d\psi = \psi_q\, dq + \psi_\vartheta\, d\vartheta = \psi_\vartheta\, d\vartheta,$$

and from the transformation equations (4)

$$dx = -\frac{\psi_\vartheta}{\varrho q} \sin\vartheta\, d\vartheta,$$

$$dy = \frac{\psi_\vartheta}{\varrho q} \cos\vartheta\, d\vartheta.$$

Hence

$$\frac{dy}{dx} = -\cot\vartheta, \quad \text{if } d\vartheta \neq 0.$$

Thus an element of any curve in the hodograph plane on which $d\vartheta \neq 0$ maps into an element with slope $-\cot\vartheta$ in the x, y-plane, i.e. normal to the streamline. Next considering the map of a curve in the hodograph plane with $d\vartheta = 0$ and using q as a parameter on the corresponding curve in the physical plane we find from the transformation equations (4) that $\frac{dx}{dq}$, $\frac{dy}{dq}$ vanish and

$$(8) \qquad \begin{cases} \dfrac{d^2x}{dq^2} = \dfrac{(\gamma-1)\cos\vartheta}{\varrho a^5}\psi_\vartheta - \dfrac{\sin\vartheta}{\varrho a}\psi_\vartheta \dfrac{d^2\vartheta}{dq^2}, \\ \dfrac{d^2y}{dq^2} = \dfrac{(\gamma-1)\sin\vartheta}{\varrho a^5}\psi_\vartheta + \dfrac{\cos\vartheta}{\varrho a}\psi_\vartheta \dfrac{d^2\vartheta}{dq^2}. \end{cases}$$

Not both of these second derivatives can vanish because the determinant of 1 and $\frac{d^2\vartheta}{dq^2}$ on the right is

$$\frac{\gamma-1}{\varrho^2 a^6}\psi_\vartheta^2 \neq 0.$$

Now a curve given parametrically by $x(q), y(q)$ will have a cusp if $\frac{dx}{dq}=\frac{dy}{dq}=0$ but not both $\frac{d^2x}{dq^2}$, $\frac{d^2y}{dq^2}$ vanish. Hence the direction $d\vartheta=0$ (i.e. the radial direction) is the exceptional direction and all curves in the hodograph plane having this direction map into cusped curves.

Curves with the non-exceptional direction are the isovel (sonic line) and the equipotential (from Chaplygin's equations $\varphi_\vartheta=0$ on the sonic line, so that the equipotential coincides with the isovel).

Curves with the exceptional direction are the streamline $\left(\text{since } \frac{d\vartheta}{dq}=-\frac{\psi_q}{\psi_\vartheta}=0\right)$ and the isocline $\left(\text{which according to (8) with } \frac{d^2\vartheta}{dq^2}=0 \text{ maps into a cusped curve with a cusp in the streamline direction}\right)$.

Since the characteristics C^+, C^- are perpendicular to the streamline and since their images Γ^+ and Γ^- are separated by the isocline in the hodograph plane they are both tangent to the sonic limit line in the physical plane and lie on opposite sides of the isocline (and hence streamline).

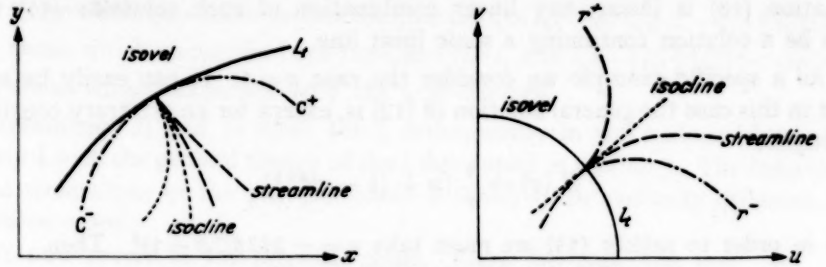


Fig. 1. Geometry in the neighborhood of the sonic limit line

Thus it is seen that the geometric behavior of curves at the sonic limit line (see fig. 1) differs from that at the L_1 and L_2. For example the isoclines are not enveloped at the L_t and no longer does one family of characteristics have cusps at the L_t. However other properties, and in particular the most important one that two branches of the solution meet at the limit line, hold true also for the L_t.

4. Construction of Solutions with Sonic Limit Line. We shall now show how flows with an L_t may be constructed. The discussion of the flow corresponding to a particular solution of the hodograph equations is facilitated by introducing the variable $\tau=q^2$ in place of q. With this variable Chaplygin's equations (1) take the form

$$(9)\qquad \begin{cases} \varphi_\vartheta = 2\tau(1-\tau)^{-\beta}\psi_\tau, \\ [1-(1+2\beta)\tau]\psi_\vartheta = -2\tau(1-\tau)^{\beta+1}\varphi_\tau, \end{cases}$$

where $\beta = \frac{1}{\gamma - 1}$. On eliminating φ we have

$$2\tau(1-\tau)^{\beta+1}\frac{\partial}{\partial\tau}\left[2\tau(1-\tau)^{-\beta}\frac{\partial\psi}{\partial\tau}\right] + [1-(1+2\beta)\tau]\frac{\partial^2\psi}{\partial\vartheta^2} = 0. \tag{10}$$

One solution of this equation is the radial (source) flow $\psi = k\vartheta$ discussed by G. I. Taylor [1]. Other solutions may be obtained by the method of separation of variables:

$$\psi_n(\tau,\vartheta) = \tau^{n/2}H_n(\tau)\begin{matrix}\sin n\vartheta \\ \cos n\vartheta,\end{matrix} \tag{11}$$

where $H_n(\tau)$ is a solution of the hypergeometric equation

$$\tau(1-\tau)H_n'' + [(n+1) - (n+1-\beta)\tau]H_n' + \tfrac{1}{2}n(n+1)\beta H_n = 0, \tag{12}$$

primes indicating differentiation with respect to τ.

For any n, a solution containing a sonic limit line can be obtained by choosing $H_n(\tau)$ to be a solution of (12) satisfying

$$\frac{d}{d\tau}[\tau^{n/2}H_n(\tau)] = 0, \quad \text{at} \quad \tau = \tau_l \tag{13}$$

where $\tau_l = 1/(2\beta+1)$. For then $\frac{\partial\psi_n}{\partial q} = 0$ for $q = 1/\sqrt{2\beta+1} = q_l$. Since the equation (10) is linear, any linear combination of such solutions (11) will also be a solution containing a sonic limit line.

As a specific example we consider the case $n = 1$. It can easily be seen that in this case the general solution of (12) is, except for an arbitrary constant multiplier

$$H_1(\tau) = \frac{1}{\tau}[\varkappa + (1-\tau)^{\beta+1}],$$

and in order to satisfy (13) we must take $\varkappa = -2(2\beta/2\beta+1)^{\beta}$. Then

$$\psi_1(\tau,\vartheta) = \tau^{-\frac{1}{2}}[\varkappa + (1-\tau)^{\beta+1}]\begin{matrix}\sin\vartheta \\ \cos\vartheta,\end{matrix} \tag{14}$$

and correspondingly we have from (9)

$$\varphi_1(\tau,\vartheta) = \tau^{-\frac{1}{2}}\{\varkappa(1-\tau)^{-\beta} + [1+(1+2\beta)\tau]\}\begin{matrix}\cos\vartheta \\ \sin\vartheta.\end{matrix} \tag{15}$$

We shall now give a discussion of the flow. Additional (non-sonic) limit singularities in the hodograph plane will occur when $\psi_\xi = 0$ or $\psi_\eta = 0$. These relations in our case (14) take the form

$$\tan(180^\circ - \vartheta) = \left[\frac{(2\beta+1)\tau - 1}{1-\tau}\right]^{\frac{1}{2}}\left\{\frac{\varkappa + (1-\tau)^{\beta+1}}{\varkappa + [1+(1+2\beta)\tau](1-\tau)^{\beta}}\right\}, \tag{16}$$

and

$$\tan\vartheta = \left[\frac{(2\beta+1)\tau - 1}{1-\tau}\right]^{\frac{1}{2}}\left\{\frac{\varkappa + (1-\tau)^{\beta+1}}{\varkappa + [1+(1+2\beta)\tau](1-\tau)^{\beta}}\right\}. \tag{17}$$

and correspond to the images l_1 and l_2 of limit lines L_1 and L_2, respectively. They meet the image l_t of the sonic limit line L_t at the point $\tau = 1/(2\beta+1)$, $\vartheta = \pi/2$. At this point $\psi_\xi = \psi_\eta = 0$ as well as $\varphi_\xi = \varphi_\eta = 0$; hence the point furnishes an example of a sonic double limit point.

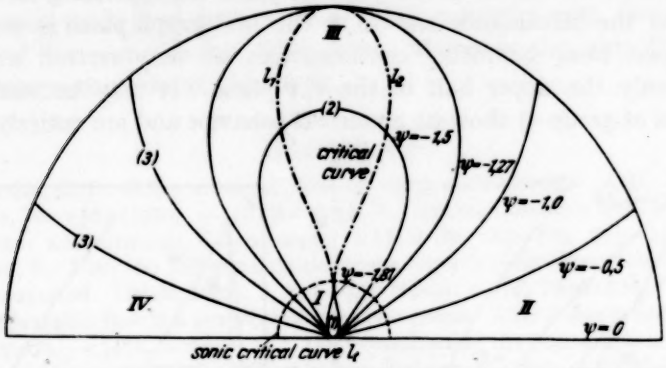


Fig. 2. Hodograph plane

A graph of the critical curves in the hodograph plane as well as selected streamlines for $\gamma = 7/5$ (corresponding to air) is presented in fig. 2. By symmetry we need only consider the upper half of the τ, ϑ-plane, i.e. $0 \leq \vartheta \leq \pi$. The streamlines may be divided into three classes:

1) those which encounter neither the l_t nor the l_1 or l_2 $(-\infty < \psi < -175\sqrt{5}/216)$.

2) those which encounter both the l_t and the l_1 and l_2 $(-175\sqrt{5}/216 < \psi < \varkappa)$.

3) those which encounter only the l_t $(\varkappa < \psi < 0)$.

All streamlines 2) and 3) meet the l_t orthogonally in the hodograph plane, in accord with the general theory of the l_t developed in section 3. The behavior of the streamlines in the physical plane is found to be radically different in the three cases.

We now revert to the physical plane. From (4), (14), and (15)

$$dx = -[\varkappa\tau^{-1}(1-\tau)^{-\beta} + \beta + \tau^{-1}]\sin 2\vartheta\, d\vartheta + \tfrac{1}{2}\beta\varkappa\tau^{-1}(1-\tau)^{-\beta-1}d\tau + \\ + \tfrac{1}{2}(2\beta+1)\tau^{-1}d\tau - \tfrac{1}{2}\tau^{-2}\{\varkappa(1-\tau)^{-\beta-1}[1-(1+\beta)\tau]+1\}\cos 2\vartheta\, d\tau,$$

$$dy = -[\varkappa\tau^{-1}(1-\tau)^{-\beta} + \beta + \tau^{-1}]\cos 2\vartheta\, d\vartheta + (\beta+1)\,d\vartheta - \\ - \tfrac{1}{2}\tau^{-2}\{\varkappa(1-\tau)^{-\beta-1}[1-(1+\beta)\tau]+1\}\sin 2\vartheta\, d\tau.$$

Hence on integrating and deleting a non-essential constant of integration

$$(18)\quad \begin{cases} 2x = [\varkappa\tau^{-1}(1-\tau)^{-\beta} + \beta + \tau^{-1}]\cos 2\vartheta + \varkappa\beta\int \tau^{-1}(1-\tau)^{-\beta-1}d\tau + \\ \qquad\qquad + (2\beta+1)\log\tau, \\ 2y = [\varkappa\tau^{-1}(1-\tau)^{-\beta} + \beta + \tau^{-1}]\sin 2\vartheta - (\beta+1)(2\vartheta). \end{cases}$$

For $\beta = 5/2$ $(\gamma = 7/5)$ the integral in the first expression of (17) can be evaluated in closed form and is in fact

$$\int \tau^{-1}(1-\tau)^{-\beta-1}d\tau = \log\frac{1-\sqrt{1-\tau}}{1+\sqrt{1-\tau}} + \frac{2}{5}\frac{1}{(1-\tau)^{\frac{5}{2}}} + \frac{2}{3}\frac{1}{(1-\tau)^{\frac{3}{2}}} + 2\frac{1}{(1-\tau)^{\frac{1}{2}}}.$$

The flow in the physical plane can now be obtained by eliminating τ and ϑ between equations (18) and (14). The limit lines are given by (18), (16), and (17).

A sketch of the curves in the physical plane corresponding to the critical curves and the streamlines shown in the hodograph plane is presented in fig. 3. Again from symmetry considerations we may restrict ourselves to mapping only the upper half of the τ, ϑ-plane. It may be seen that the streamlines of group 1) show no abnormal behavior and are entirely subsonic;

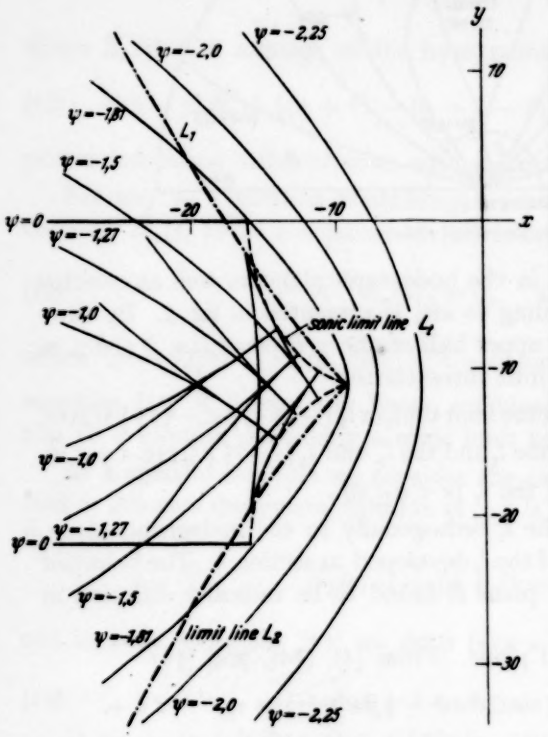


Fig. 3. Physical plane

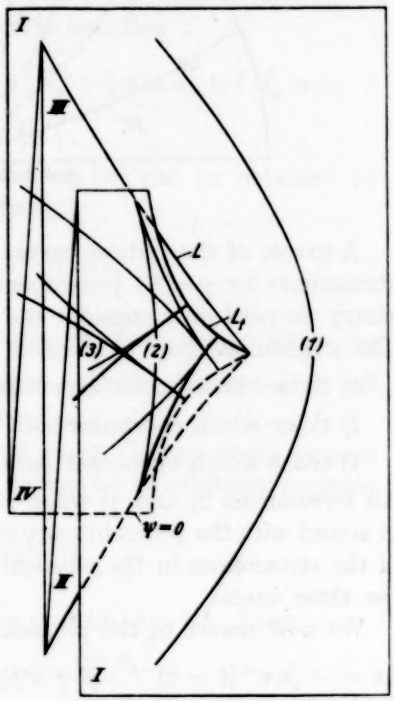


Fig. 4. Sketch of multi-sheeted surface corresponding to physical plane

however, those of group 2) and 3) have singular points on the limit lines. More specifically, the streamlines of group 3) exhibit a cusp on meeting a branch of the L_t and those of group 2) meet the L_1 and L_2 and each branch of the L_t in a cusp. The behavior of the streamlines of group 2) and 3) becomes clearer if the physical plane is envisaged as a quadruply-sheeted surface (see fig. 4). The individual sheets correspond to the regions (denoted by Roman numerals in fig. 2) of the hodograph plane which are bounded by the arcs of the critical curves and the straight streamlines. On each reflection at a limit line the streamline must then be thought of as passing from one sheet to another. In this example there are no regular streamlines along which the flow is supersonic as well as subsonic. Only the streamline

which separates the streamlines of groups 1) and 2) reaches sonic speed on passing through the cusp of the L_t.

The isovels ($\tau = \text{const}$) in the physical plane are portions of trochoids; in the special case $\tau = 1/(2\beta + 1)$ of the L_t, the curve is a segment of a cycloid. The non-sonic limit lines form a cusp which coincides with the cusp point of the L_t at the sonic double limit point. Moreover the cusp directions of the two curves coincide at this point.

Bibliography

[1] TAYLOR, G. I.: Some cases of flow of compressible fluids. ARC Reports and Memoranda, No. 1382 (1930). — [2] RINGLEB, F.: Exakte Lösungen der Differentialgleichungen einer adiabatischen Gasströmung. ZAMM **20**, 185—198, 295—296 (1940). — [3] RINGLEB, F.: Über die Differentialgleichungen einer adiabatischen Gasströmung und den Strömungsstoß. Dtsch. Math. **5**, 377—384 (1940). — [4] TEMPLE, G., and J. YARWOOD: Compressible flow in a convergent-divergent nozzle. ARC Reports and Memoranda, No. 2077 (1942). — [5] CRAGGS, J. W.: The breakdown of the hodograph transformation for irrotational compressible fluid flow in two dimensions. Proc. Cambridge Phil. Soc. **44**, 360—379 (1948). — [6] LIGHTHILL, M. J.: The hodograph method. Chap. VII in Modern Developments in Fluid Dynamics. High Speed Flow. (Ed. by L. Howarth). Oxford: 1953. — [7] GEIRINGER, H.: Grenzlinien der Hodographentransformation. Math. Z. **63**, 514—524 (1956).

Department of Mathematics, University of Maryland, College Park, Maryland

(Eingegangen am 1. Dezember 1956)

SCHÄFKE, F. W.
Math. Zeitschr. Bd. 67, S. 238—251 (1957)

Integrale über Produkte von Sphäroidfunktionen

Von

FRIEDRICH WILHELM SCHÄFKE

Einleitung

Die vorliegende Note beschäftigt sich mit Sphäroidfunktionen, d.h. mit den Lösungen der Differentialgleichung

$$[(1-z^2)\,y'(z)]' + \left[\lambda + \gamma^2(1-z^2) - \frac{\mu^2}{1-z^2}\right] y(z) = 0,$$

die bei der Separation der Schwingungsgleichung $\Delta u + k^2 u = 0$ in rotationselliptischen Koordinaten auftritt. Die Note schließt sich eng an das Werk von J. MEIXNER und F. W. SCHÄFKE [*1*][1]) an, zu dem sie in verschiedener Hinsicht Ergänzungen gibt.

Betrachtet werden zwei Typen von Integralrelationen, die Produkte von Sphäroidfunktionen betreffen.

Den im Abschnitt 1. behandelten Integralrelationen 1. Art liegt der Gedanke zugrunde, die von den Sturm-Liouvilleschen Eigenwertproblemen her vertraute Methode zum Beweise von Orthogonalitätsrelationen zu verknüpfen mit Aussagen über das asymptotische Verhalten von Sphäroidfunktionen für große Werte des Arguments. Man betrachtet zwei Funktionen $y_1(z)$, $y_2(z)$ zum gleichen γ^2, jedoch zu verschiedenen Parameterpaaren λ_1, μ_1^2 bzw. λ_2, μ_2^2. Schreibt man die entsprechenden Differentialgleichungen auf, multipliziert mit y_2 bzw. y_1, subtrahiert und integriert längs eines Weges $\mathfrak{C}$ zwischen a und b, der $+1$ und -1 nicht treffen möge, so entsteht

$$(1)\quad \begin{cases} (z^2-1)\left(y_1'(z)\,y_2(z) - y_1(z)\,y_2'(z)\right)\Big|_{a(\mathfrak{C})}^{b} \\ \qquad = (\lambda_1 - \lambda_2)_{(\mathfrak{C})} \int_a^b y_1(z)\,y_2(z)\,dz + (\mu_1^2 - \mu_2^2)_{(\mathfrak{C})} \int_a^b (z^2-1)^{-1}\,y_1(z)\,y_2(z)\,dz. \end{cases}$$

Hier ist es nun möglich, für Funktionen y_1, y_2 mit geeignetem asymptotischen Verhalten bzw. Umlaufsverhalten bzw. mit geeigneten Eigenschaften bei $+1$ (oder -1) a und b längs gewisser Strahlen gegen ∞ rücken zu lassen oder auch etwa $a \to +1$, derart daß dabei der Grenzwert der linken Seite explizit angegeben werden kann und die Existenz der Integrale rechts feststeht. Diese Methode dürfte sich zuerst bei FLAMMER [*2*] finden (vgl. das Zitat bei I. MARX [*3*]), jedoch dort nur für $\mu_1^2 = \mu_2^2$, für Funktionen ganzer Ordnung und ganzen Grades und für $a \to +1$. Die in 1. gegebenen Relationen stellen

[1]) Ziffern in eckigen Klammern verweisen auf die am Schluß notierte Literatur.

wesentliche Verallgemeinerungen dar. Als bemerkenswerte Folgerung ergibt sich eine Formel für die Ableitung $\frac{\partial \lambda}{\partial \nu}$.

Die in 2. behandelten Integralrelationen 2. Art knüpfen an die Bemerkung an, daß mit einer Funktion u auch deren Ableitungen nach einer kartesischen Koordinate der Schwingungsgleichung genügen, also gemäß [1], Abschnitt 1.133., Satz 1 zur Aufstellung von Kernen für Integralbeziehungen zwischen Sphäroidfunktionen dienen können. Wählt man hier u als in rotationselliptischen Koordinaten separierte Lösung der Schwingungsgleichung, so erhält man gewisse Integrale, die Produkte von Sphäroidfunktionen enthalten. Dieser Gedanke wird hier für beliebig komplexe Parameter und Variable vollständig durchgeführt. Die Methode findet sich wohl zuerst bei WHITTAKER [5] für Mathieusche Funktionen, für dieselben Funktionen allgemeiner in [1], Abschnitt 2.83. Einen entsprechenden Ansatz für Sphäroidfunktionen, jedoch nur für Funktionen ganzer Ordnung und ganzen Grades im Reellen, hat MARX [3] gegeben. Allerdings fehlt dort die explizite Angabe der Relationen, da die entscheidenden Konstanten nicht in geeigneter Weise bestimmt werden. Ferner finden sich dort etwas irreführende Bemerkungen über die Auffassung derartiger Integralrelationen als „Rekursionsformeln" für Sphäroidfunktionen.

Hierzu, d.h. zur Frage der Existenz derartiger Rekursionen, finden sich in 4. einige kritische Bemerkungen.

Zuvor wird in 3. kurz gezeigt, was die angegebenen Methoden für andere Klassen von speziellen Funktionen liefern. Insbesondere wird die zweite genannte Methode für Zylinderfunktionen und Kugelfunktionen angesetzt. Es zeigt sich, daß man hier in sehr einfacher Weise die entsprechenden Rekursionsformeln erhält. Das liegt daran, daß hier die entstehenden Kerne in einfacher Weise ausgeartet sind. Das ist jedoch für Mathieusche Funktionen und Sphäroidfunktionen nicht mehr der Fall.

1. Integralrelationen 1. Art

11. Die Grundrelation

Wir knüpfen an die in der Einleitung gegebene Formel (1) an. Es seien ν_1 bzw. ν_2 charakteristische Exponenten zu $\lambda_1, \gamma^2, \mu_1^2$ bzw. zu $\lambda_2, \gamma^2, \mu_2^2$ mit $\gamma^2 \neq 0$, derart daß

$$\lambda_1 = \lambda_{\nu_1}^{\mu_1}(\gamma^2), \qquad \lambda_2 = \lambda_{\nu_2}^{\mu_2}(\gamma^2) \tag{2}$$

(vgl. [1], 3.531.). Dabei sei *vorausgesetzt*

$$\nu_1 \not\equiv \tfrac{1}{2} \pmod 1, \qquad \nu_2 \not\equiv \tfrac{1}{2} \pmod 1.$$

Wir wählen nun (vgl. [1], 3.64.)

$$y_1(z) = S_{\nu_1}^{\mu_1(1)}(z;\gamma), \qquad y_2(z) = S_{\nu_2}^{\mu_2(1)}(z;\gamma). \tag{3}$$

Dabei ist das Vorzeichen von γ willkürlich. Wir setzen nach seiner Festlegung

$$\alpha = -\arg\gamma \tag{4}$$

mit der Verabredung

$$-\pi < \alpha \leq \pi. \tag{5}$$

Dann gilt einmal für den Umlauf um $+1$ und -1 zugleich ([1], 3.65. (48))

$$y_1(z\,e^{2\pi i}) = e^{2\pi i\nu_1} y_1(z), \quad y_2(z\,e^{2\pi i}) = e^{2\pi i\nu_2} y_2(z), \tag{6}$$

andererseits nach [1], 3.65. (49), (50) für den Grenzübergang

$$z \to \infty, \quad \arg z = \alpha \qquad (\arg \gamma z = 0)$$

sicher ($\varkappa = 1, 2$)

$$\begin{cases} y_\varkappa(z) = \dfrac{1}{\gamma z}\cos\left(\gamma z - \dfrac{\nu_\varkappa + 1}{2}\pi\right) + O\left(\dfrac{1}{z^2}\right), \\ y_\varkappa'(z) = -\dfrac{1}{z}\sin\left(\gamma z - \dfrac{\nu_\varkappa + 1}{2}\pi\right) + O\left(\dfrac{1}{z^2}\right) \end{cases} \tag{7}$$

(zur letzten Zeile vgl. [1], 3.64., S. 291). Daraus folgt, daß beim gleichen Grenzübergang gilt

$$(z^2 - 1)\,(y_1'(z)\,y_2(z) - y_1(z)\,y_2'(z)) \to \frac{1}{\gamma}\sin\frac{\nu_1 - \nu_2}{2}\pi. \tag{8}$$

Wir wählen nun in (1) den Integrationsweg der Fig. 1.

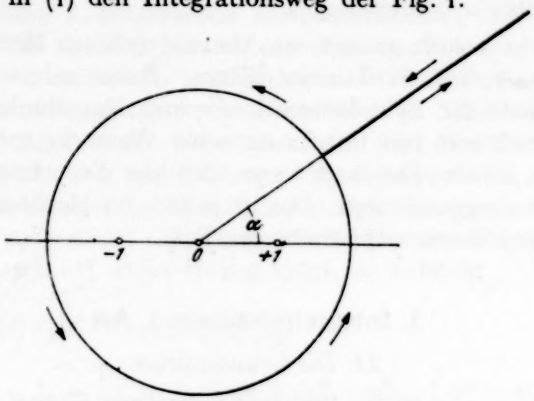


Fig. 1. Der Integrationsweg für (9)

Dann hat man mit Hilfe von (6) sofort[2])

$$\begin{cases} \left(\lambda_{\nu_1}^{\mu_1}(\gamma^2) - \lambda_{\nu_2}^{\mu_2}(\gamma^2)\right) \overset{(+1+,\,-1+)}{\underset{\infty\, e^{i\alpha}}{\int}} S_{\nu_1}^{\mu_1(1)}(z;\gamma)\, S_{\nu_2}^{\mu_2(1)}(z;\gamma)\, dz + \\ \quad + (\mu_1^2 - \mu_2^2) \overset{(+1+,\,-1+)}{\underset{\infty\, e^{i\alpha}}{\int}} (z^2-1)^{-1} S_{\nu_1}^{\mu_1(1)}(z;\gamma)\, S_{\nu_2}^{\mu_2(1)}(z;\gamma)\, dz \\ \qquad = (e^{2\pi i(\nu_1+\nu_2)} - 1)\,\dfrac{1}{\gamma}\sin\dfrac{\nu_1 - \nu_2}{2}\pi. \end{cases} \tag{9}$$

[2]) Zur Bezeichnung vgl. [1], S. 251 ff.

Es zeigt sich nun, daß die Wahl (3) keine besondere Einschränkung darstellt. Denn man hat stets

$$\lambda_\nu^\mu(\gamma^2) = \lambda_{-\nu-1}^\mu(\gamma^2),$$

und $S_\nu^{\mu(1)}(z;\gamma)$, $S_{-\nu-1}^{\mu(1)}(z;\gamma)$ bilden ein Fundamentalsystem. Man kann daher ν_1 durch $-\nu_1-1$ oder bzw. und ν_2 durch $-\nu_2-1$ ersetzen. Man hat dann vier Formeln der Form (9), mit deren Hilfe alle Ausdrücke

$$(\lambda_1-\lambda_2)\int y_1 y_2\,dz + (\mu_1^2-\mu_2^2)\int (z^2-1)^{-1} y_1 y_2\,dz$$

ausgewertet werden können, falls für die $y_\varkappa(z)$ $(\varkappa=1,2)$ die Darstellungen als Linearkombinationen von $S_{\nu_\varkappa}^{\mu_\varkappa(1)}(z;\gamma)$, $S_{\nu_\varkappa-1}^{\mu_\varkappa(1)}(z;\gamma)$ gegeben sind. Wir gehen hier auf derartige Folgerungen allgemein nicht näher ein.

Einfach werden natürlich die Formeln für $\lambda_1=\lambda_2$ oder $\mu_1^2=\mu_2^2$. Hier sei nur noch einiges für den zweiten Fall besprochen. Wir setzen im folgenden

$$\mu_1=\mu_2=\mu.$$

Unter dieser Annahme sollen einige Spezialfälle und Ausartungen von (9) notiert werden.

12. Integrationswege mit Umlauf um +1

Zuerst eine grundsätzliche Überlegung.

Wir wählen y_1 und y_2 so, daß ihr Produkt bei $+1$ regulär analytisch wird, also etwa (γ^2 sei nicht Ausnahmewert)

$$y_1 = \mathrm{Ps}_{\nu_1}^{\mu}(z;\gamma^2), \qquad y_2 = \mathrm{Ps}_{\nu_2}^{-\mu}(z;\gamma^2).$$

Dann kann der Integrationsweg über $+1$ hinweggezogen werden. Nach Substitution $z\to -z$ erhält man dann entsprechende Integrale mit dem Integrationsweg der Fig. 2. Diese Herleitung ist wegen der doppelten Umrechnung vielleicht etwas umständlich. Es kam nur darauf an, zu zeigen, daß die im folgenden gegebenen Relationen als Folgerungen aus (9) aufgefaßt werden können.

Wir wählen zur Aufstellung der Relationen mit dem Integrationsweg der Fig. 2 lieber einen mehr direkten Weg.

Wir nehmen für

$$\alpha = -\arg\gamma$$

jetzt (wegen der Stelle -1)

$$|\alpha| < \pi \tag{10}$$

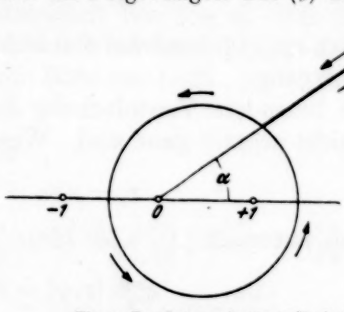


Fig. 2. Der Integrationsweg für (14)

an, was wegen $\gamma\to-\gamma$ keine Einschränkung bedeutet. Wir wählen dann

$$\begin{cases} \lambda_1=\lambda_{\nu_1}^{\mu}(\gamma^2), & \lambda_2=\lambda_{\nu_2}^{\mu}(\gamma^2), & \nu_\varkappa \not\equiv \tfrac{1}{2} \pmod 1, \\ y_1=\mathrm{Ps}_{\nu_1}^{\mu}(z;\gamma^2), & y_2=\mathrm{Ps}_{\nu_2}^{\mu}(z;\gamma^2). \end{cases} \tag{11}$$

Dann multipliziert sich $y_1(z) \cdot y_2(z)$ bei positivem Umlauf um $+1$ mit $e^{-2\pi i\mu}$. Da wir auch das asymptotische Verhalten benötigen, notieren wir die Formel (nach [1], 3.61. (4) und 3.66. (56))

$$(12)\quad \begin{cases} \mathrm{Ps}_\nu^\mu(z;\gamma^2) \\ = \dfrac{1}{\cos\nu\pi}\left[\dfrac{1}{\Gamma(\nu-\mu+1)\,V_\nu^\mu(\gamma)}\,S_\nu^{\mu\,(1)}(z;\gamma) - \dfrac{1}{\Gamma(-\nu-\mu)\,V_{-\nu-1}^\mu(\gamma)}\,S_{-\nu-1}^{\mu\,(1)}(z;\gamma)\right]. \end{cases}$$

Mit ihrer Hilfe kann für den Grenzübergang

$$z\to\infty, \qquad \arg z=\alpha \qquad (\arg\gamma z=0)$$

erhalten werden:

$$(13)\quad \begin{cases} (z^2-1)\,(y_1'(z)\,y_2(z)-y_1(z)\,y_2'(z)) \to \dfrac{1}{\gamma\cos\nu_1\pi\cos\nu_2\pi}\times \\ \quad\times\Big\{\sin\dfrac{\nu_1-\nu_2}{2}\pi\cdot\Big[\dfrac{1}{\Gamma(\nu_1-\mu+1)\,\Gamma(\nu_2-\mu+1)\,V_{\nu_1}^\mu(\gamma)\,V_{\nu_2}^\mu(\gamma)} - \\ \qquad - \dfrac{1}{\Gamma(-\nu_1-\mu)\,\Gamma(-\nu_2-\mu)\,V_{-\nu_1-1}^\mu(\gamma)\,V_{-\nu_2-1}^\mu(\gamma)}\Big] + \\ \quad + \sin\dfrac{\nu_1+\nu_2+1}{2}\pi\cdot\Big[\dfrac{1}{\Gamma(-\nu_1-\mu)\,\Gamma(\nu_2-\mu+1)\,V_{-\nu_1-1}^\mu(\gamma)\,V_{\nu_2}^\mu(\gamma)} - \\ \qquad - \dfrac{1}{\Gamma(\nu_1-\mu+1)\,\Gamma(-\nu_2-\mu)\,V_{\nu_1}^\mu(\gamma)\,V_{-\nu_2-1}^\mu(\gamma)}\Big]\Big\}. \end{cases}$$

Wir kürzen die rechte Seite mit $C(\nu_1,\nu_2,\mu,\gamma^2)$ ab. Aus (1) folgt dann mit (13) und der Umlaufseigenschaft:

$$(14)\quad \left(\lambda_{\nu_1}^\mu(\gamma^2)-\lambda_{\nu_2}^\mu(\gamma^2)\right)\int_{\infty e^{i\alpha}}^{(+1+)} \mathrm{Ps}_{\nu_1}^\mu(z;\gamma^2)\,\mathrm{Ps}_{\nu_2}^\mu(z;\gamma^2)\,dz = (e^{-2\pi i\mu}-1)\,C(\nu_1,\nu_2,\mu,\gamma^2).$$

13. Integrationsstrahl von $+1$ aus

Ist $\Re\,\mu<1$, so ist Integration an $+1$ heran möglich und man hat aus (14):

$$(15)\quad \left(\lambda_{\nu_1}^\mu(\gamma^2)-\lambda_{\nu_2}^\mu(\gamma^2)\right)\int_1^{\infty e^{i\alpha}} \mathrm{Ps}_{\nu_1}^\mu(z;\gamma^2)\,\mathrm{Ps}_{\nu_2}^\mu(z;\gamma^2)\,dz = C(\nu_1,\nu_2,\mu,\gamma^2).$$

Dies folgt aus (14) zunächst für nicht ganzes μ, damit dann allgemein durch Grenzübergang.

Eine besondere Vereinfachung ergibt sich für den Fall, daß $\nu_1+\mu$ und $\nu_2+\mu$ nicht-negativ ganz sind. Wegen $\Re\,\mu<1$ ist dann von selbst

$$\nu_\varkappa-\mu \neq -1,-2,-3,\ldots$$

und man hat nach [1], 3.66. (67):

$$S_{\nu_\varkappa}^{\mu\,(1)}(z;\gamma) = K_{\nu_\varkappa}^\mu(\gamma)\,\mathrm{Ps}_{\nu_\varkappa}^\mu(z;\gamma^2).$$

Damit kann man gemäß (13), (15) — oder auch leicht direkt mit (8) — schreiben:

$$(16)\quad \left(\lambda_{\nu_1}^\mu(\gamma^2)-\lambda_{\nu_2}^\mu(\gamma^2)\right)\int_1^{\infty e^{i\alpha}} S_{\nu_1}^{\mu\,(1)}(z;\gamma)\,S_{\nu_2}^{\mu\,(1)}(z;\gamma)\,dz = \frac{1}{\gamma}\sin\frac{\nu_1-\nu_2}{2}\pi.$$

Für ganze μ und damit auch ganze ν_1, ν_2, ferner für $\gamma > 0$ entsteht daraus die Formel von FLAMMER [2], bzw. von MARX [3], S. 271, Formel (7).

14. Eine Folgerung für die Ableitung $\frac{\partial \lambda}{\partial \nu}$.

Wir knüpfen an (9) an, wo wir wieder $\mu_1 = \mu_2 = \mu$ wählen. Wir wollen voraussetzen, daß γ^2 nicht Ausnahmewert für $\lambda_\nu^\mu(\gamma^2)$ ist (vgl. [1], 3.531.). Wir setzen $\nu_2 = -\nu - 1$ und machen den Grenzübergang $\nu_1 \to \nu$, nachdem wir beide Seiten durch $\nu_1 - \nu$ dividiert haben. Im Integral kann offenbar einfach $\nu_1 = \nu$ gesetzt werden und man kann mit [1], 3.66. (56), 3.544. (26) auswerten:

$$\begin{aligned}
&\int_{\infty e^{i\alpha}}^{(+1+,\,-1+)} S_\nu^{\mu(1)}(z;\gamma)\, S_{-\nu-1}^{\mu(1)}(z;\gamma)\, dz \\
&\qquad = V_\nu^\mu(\gamma)\, V_{-\nu-1}^\mu(\gamma) \oint \tilde{Q}s_\nu^\mu(z;\gamma^2)\, \tilde{Q}s_{-\nu-1}^\mu(z;\gamma^2)\, dz \\
&\qquad = 2\pi i \frac{\cos \nu\pi}{2\nu+1} V_\nu^\mu(\gamma)\, V_{-\nu-1}^\mu(\gamma).
\end{aligned}$$

Berücksichtigt man noch [1], 3.66. (58), so hat man schließlich die vielleicht bemerkenswerte Relation

$$\frac{\partial \lambda_\nu^\mu(\gamma^2)}{\partial \nu} = (2\nu+1)\, A_\nu^\mu(\gamma^2)\, A_\nu^{-\mu}(\gamma^2). \tag{17}$$

Man kann sie natürlich für $\gamma^2 = 0$, $\lambda_\nu^\mu = \nu(\nu+1)$ unmittelbar bestätigen.

2. Integralrelationen 2. Art

21. Vorbemerkungen

Wir knüpfen an den Abschnitt über gestreckt-rotationselliptische Koordinaten 1.123. in [1] an. Wir betrachten eine Funktion u einmal in Abhängigkeit von x_1, x_2, x_3, zum andern in Abhängigkeit von ξ, η, φ. Wir drücken die partiellen Ableitungen nach x_1, x_2, x_3 durch die nach ξ, η, φ aus. Das kann entweder ziemlich direkt oder auch mit Hilfe von 1.123., (15*) geschehen. Wir notieren:

$$\left\{\begin{aligned}
c\frac{\partial u}{\partial x_3} &= \frac{\eta(\xi^2-1)}{\xi^2-\eta^2}\frac{\partial u}{\partial \xi} + \frac{\xi(1-\eta^2)}{\xi^2-\eta^2}\frac{\partial u}{\partial \eta}, \\
c\frac{\partial u}{\partial x_1} &= \frac{\xi\sqrt{(\xi^2-1)(1-\eta^2)}}{\xi^2-\eta^2}\cos\varphi\frac{\partial u}{\partial \xi} - \frac{\eta\sqrt{(\xi^2-1)(1-\eta^2)}}{\xi^2-\eta^2}\cos\varphi\frac{\partial u}{\partial \eta} - \\
&\qquad - \frac{\sin\varphi}{\sqrt{(\xi^2-1)(1-\eta^2)}}\frac{\partial u}{\partial \varphi}, \\
c\frac{\partial u}{\partial x_2} &= \frac{\xi\sqrt{(\xi^2-1)(1-\eta^2)}}{\xi^2-\eta^2}\sin\varphi\frac{\partial u}{\partial \xi} - \frac{\eta\sqrt{(\xi^2-1)(1-\eta^2)}}{\xi^2-\eta^2}\sin\varphi\frac{\partial u}{\partial \eta} + \\
&\qquad + \frac{\cos\varphi}{\sqrt{(\xi^2-1)(1-\eta^2)}}\frac{\partial u}{\partial \varphi}.
\end{aligned}\right. \tag{18}$$

Wir fassen noch zusammen:

$$(19)\quad \begin{cases} c\left(\dfrac{\partial u}{\partial x_1} \pm i \dfrac{\partial u}{\partial x_2}\right) \\ \qquad = e^{\pm i\varphi}\left\{\dfrac{\sqrt{(\xi^2-1)(1-\eta^2)}}{\xi^2-\eta^2}\left[\xi\dfrac{\partial u}{\partial \xi} - \eta\dfrac{\partial u}{\partial \eta}\right] \pm i\dfrac{1}{\sqrt{(\xi^2-1)(1-\eta^2)}}\dfrac{\partial u}{\partial \varphi}\right\}. \end{cases}$$

Nun bemerken wir:

Ist $u(\xi, \eta, \varphi)$ eine für ξ, η, φ in einem Gebiete bezüglich der drei komplexen Variablen, in dem $\xi \neq \pm 1$, $\eta \neq \pm 1$, $\xi \neq \pm \eta$ gilt, regulär analytische Lösung der Schwingungsgleichung in gestreckt-rotationselliptischen Koordinaten ($\gamma = kc$)

$$(20)\quad \frac{\partial}{\partial \xi}\left((\xi^2-1)\frac{\partial u}{\partial \xi}\right) + \frac{\partial}{\partial \eta}\left((1-\eta^2)\frac{\partial u}{\partial \eta}\right) + \frac{\xi^2-\eta^2}{(\xi^2-1)(1-\eta^2)}\frac{\partial^2 u}{\partial \varphi^2} + \gamma^2(\xi^2-\eta^2)u = 0,$$

so gilt dasselbe für die mit u gebildeten Ausdrücke auf den rechten Seiten von (18) und (19).

Das folgt aus der Tatsache, daß mit u auch $\dfrac{\partial u}{\partial x_\varkappa}$ ($\varkappa = 1, 2, 3$) der Gleichung

$$\frac{\partial^2 u}{\partial x_1^2} + \frac{\partial^2 u}{\partial x_2^2} + \frac{\partial^2 u}{\partial x_3^2} + k^2 u = 0$$

genügen, und aus der Bemerkung, daß die Umrechnungen von x_1, x_2, x_3 auf ξ, η, φ im Komplexen genau so verlaufen wie im Reellen.

22. Aufstellung der Kerne

Wir wählen nun speziell u als separierte Lösung von (20), also etwa ($\gamma^2 \neq 0$)

$$(21)\qquad u(\xi, \eta, \varphi) = S_\nu^{\mu(j)}(\xi; \gamma)\, \tilde{Q}s_\nu^\mu(\eta; \gamma^2)\, e^{i\mu\varphi}.$$

Mit dieser Form erfassen wir alle separierten Lösungen durch Bildung von Linearkombinationen, indem wir die möglichen Substitutionen $\mu \to -\mu$ oder bzw. und $\nu \to -\nu - 1$ und etwa $j = 3, 4$ beachten. Dabei ist stets ([*I*]. 3.65. (45), 3.543. (23*))

$$S_\nu^{-\mu(j)} = S_\nu^{\mu(j)}, \quad \tilde{Q}s_\nu^{-\mu} = \tilde{Q}s_\nu^\mu.$$

Bilden wir nun mit (21) den Ausdruck der rechten Seite der ersten Zeile von (18), so entsteht

$$(22)\qquad v_1(\xi, \eta)\, e^{i\mu\varphi}$$

mit

$$(22^*)\quad \begin{cases} v_1(\xi, \eta) = \dfrac{1}{\xi^2-\eta^2} \times \\ \qquad \times\left[\eta(\xi^2-1)\, S_\nu^{\mu(j)'}(\xi; \gamma)\, \tilde{Q}s_\nu^\mu(\eta; \gamma^2) + \xi(1-\eta^2)\, S_\nu^{\mu(j)}(\xi; \gamma)\, \tilde{Q}s_\nu^{\mu'}(\eta; \gamma^2)\right]. \end{cases}$$

Hier bedeutet, wie stets auch im folgenden, der Strich (') die Ableitung nach dem ersten Argument (ξ bzw. η).

Bilden wir mit (21) bei oberem Vorzeichen den Ausdruck auf der rechten Seite von (19), so entsteht (bis auf konstanten Faktor)

$$v_2(\xi,\eta)\, e^{i(\mu+1)\varphi} \tag{23}$$

mit

$$\begin{cases} v_2(\xi,\eta) \\ \quad = \dfrac{(\xi^2-1)^{\frac{1}{2}}(\eta^2-1)^{\frac{1}{2}}}{\xi^2-\eta^2}\left[\xi\, S_\nu^{\mu(j)\prime}(\xi;\gamma)\,\tilde{\mathrm{Q}}\mathrm{s}_\nu^\mu(\eta;\gamma^2) - \eta\, S_\nu^{\mu(j)}(\xi;\gamma)\,\tilde{\mathrm{Q}}\mathrm{s}_\nu^{\mu\prime}(\eta;\gamma^2)\right] - \\ \quad - \dfrac{\mu}{(\xi^2-1)^{\frac{1}{2}}(\eta^2-1)^{\frac{1}{2}}}\, S_\nu^{\mu(j)}(\xi;\gamma)\,\tilde{\mathrm{Q}}\mathrm{s}_\nu^\mu(\eta;\gamma^2). \end{cases} \tag{23*}$$

Bildet man dasselbe mit unterem Vorzeichen, so erhält man unten nichts Neues, da dasselbe auch mit den Substitutionen $\mu \to -\mu$ bzw. $\mu+1 \to -\mu-1$ erreicht werden kann, die unabhängig geschehen können, wie die folgenden Überlegungen sofort zeigen.

(22) und (23) sind nun nach den vorangehenden Bemerkungen Lösungen von (20), und zwar in der Form, daß die φ-Abhängigkeit schon absepariert ist. Damit sind wir in der Lage, (22*) und (23*) als Kerne gemäß [*1*], 1.133., Satz 1 zu verwenden. Dort hat man nur bei Betrachtung von (23*) μ durch $\mu+1$ zu ersetzen.

23. Die Integralrelationen

Für die Anwendung des genannten Satzes kommt es nun auf die Wahl geeigneter Integrationswege und die Erfüllung der weiteren Voraussetzungen an.

Wir wählen einen Integrationsweg der η-Ebene, der $+1$ und -1 einmal positiv umläuft, und als Funktion u_2 mit ganzem p bei dem Kern (22*) die Funktion

$$u_2(\eta) = \tilde{\mathrm{Q}}\mathrm{s}_{-\nu-p-1}^\mu(\eta;\gamma^2),$$

bei (23*) entsprechend

$$u_2(\eta) = \tilde{\mathrm{Q}}\mathrm{s}_{-\nu-p-1}^{\mu+1}(\eta;\gamma^2).$$

ξ lassen wir außerhalb dieses Integrationsweges variieren und verabreden eindeutig

$$(\xi^2-1)^{\frac{1}{2}} \sim \xi \qquad (\xi \to \infty).$$

Entsprechend wird $(\eta^2-1)^{\frac{1}{2}}$ in der längs der Strecke zwischen $+1$ und -1 aufgeschnittenen Ebene eindeutig erklärt. Dann sind offenbar alle Voraussetzungen von [*1*], 1.133., Satz 1 erfüllt auf Grund der Umlaufsrelationen

$$\tilde{\mathrm{Q}}\mathrm{s}_\nu^\mu(z\,e^{2\pi i};\gamma^2) = e^{-2\pi i(\nu+1)}\,\tilde{\mathrm{Q}}\mathrm{s}_\nu^\mu(z;\gamma^2).$$

Somit *ist*

$$\oint v_1(\xi,\eta)\,\tilde{\mathrm{Q}}\mathrm{s}_{-\nu-p-1}^\mu(\eta;\gamma^2)\,d\eta \tag{24}$$

eine Sphäroidfunktion zu den Parametern $\lambda_{\nu+p}^\mu(\gamma^2)$, γ^2, μ^2 und

$$\oint v_2(\xi,\eta)\,\tilde{\mathrm{Q}}\mathrm{s}_{-\nu-p-1}^{\mu+1}(\eta;\gamma^2)\,d\eta \tag{25}$$

eine Sphäroidfunktion zu $\lambda_{\nu+p}^{\mu+1}(\gamma^2), \gamma^2, (\mu+1)^2$. Natürlich nehmen wir an, daß γ^2 niemals Ausnahmewert ist.

Zur Entscheidung darüber, um welche Sphäroidfunktion es sich in (24) bzw. (25) handelt, untersuchen wir ihr asymptotisches Verhalten für $\xi \to \infty$ bei $|\arg \gamma\xi| \leq \pi - \delta$, $\delta > 0$, wobei wir [1], 3.65. (49), (50) beachten. Wir erkennen, daß gleichmäßig bezüglich η auf dem Integrationswege

$$(26) \qquad v_1(\xi,\eta) = \pm i \frac{1}{\xi} e^{\pm i\left(\gamma\xi - \frac{\nu+1}{2}\pi\right)}\left(\eta\, \tilde{\mathrm{Qs}}_\nu^\mu(\eta;\gamma^2) + O\left(\frac{1}{\xi}\right)\right)$$

und

$$(27) \qquad v_2(\xi,\eta) = \pm i \frac{1}{\xi} e^{\pm i\left(\gamma\xi - \frac{\nu+1}{2}\pi\right)}\left((\eta^2-1)^{\frac{1}{2}}\, \tilde{\mathrm{Qs}}_\nu^\mu(\eta;\gamma^2) + O\left(\frac{1}{\xi}\right)\right)$$

gilt, wobei das obere Vorzeichen für $j=3$, das untere für $j=4$ zu wählen ist. Stets gibt das Glied mit $S_\nu^{\mu(j)'}(\xi;\gamma)$ den Ausschlag. Da die kritische Gerade $\arg\gamma\xi = 0$ mit umfaßt wird, ist somit eindeutig durch (24) die Funktion

$$(28) \qquad \gamma(\pm i)^{p+1} \oint \eta\, \tilde{\mathrm{Qs}}_\nu^\mu(\eta;\gamma^2)\, \tilde{\mathrm{Qs}}_{-\nu-p-1}^\mu(\eta;\gamma^2)\, d\eta \cdot S_{\nu+p}^{\mu(j)}(\xi;\gamma)$$

und durch (25) die Funktion

$$(29) \qquad \gamma(\pm i)^{p+1} \oint (\eta^2-1)^{\frac{1}{2}}\, \tilde{\mathrm{Qs}}_\nu^\mu(\eta;\gamma^2)\, \tilde{\mathrm{Qs}}_{-\nu-p-1}^\mu(\eta;\gamma^2)\, d\eta \cdot S_{\nu+p}^{\mu+1(j)}(\xi;\gamma)$$

dargestellt.

Wir verzichten darauf, die erhaltenen Integralrelationen durch Zusammenfügung von (22*), (24), (28) bzw. (23*), (25), (29) bei nochmaliger Nennung der Voraussetzungen in zwei Formelzeilen aufzuschreiben.

24. Bemerkungen

a) Auf Grund der Halbumlaufsrelationen der Funktionen $\tilde{\mathrm{Qs}}$ sind die Integrale in (28), (29) beide 0, wenn p eine gerade Zahl ist. Aus demselben Grunde könnte man für ungerades p auf Integrale über einen halben Umlauf übergehen. Dort kann $(\pm i)^{p+1} = (-1)^{\frac{p+1}{2}}$ gesetzt werden.

b) Für $\nu+\mu = -1, -2, -3, \ldots$ und $-\nu-p-1-\mu = -1, -2, -3, \ldots$ bzw. $-\nu-p-2-\mu = -1, -2, -3, \ldots$ ist der Integrand in (28) bzw. (29) bei $+1$ und -1 regulär analytisch und die Integrale verschwinden ebenfalls. (Vgl. [1], 3.534. und 3.61. (4).)

c) Für $\Re\mu > -1$, $\nu+\mu = -1, -2, -3, \ldots$ und $-\nu-p-1-\mu = 0, 1, 2, \ldots$ bzw. $-\nu-p-2-\mu = 0, 1, 2, \ldots$ kann das Integral in (28) bzw. (29) — man beachte dazu Bemerkung a) — in ein Integral über die Strecke von -1 bis $+1$ umgerechnet werden, in dessen Integrand die Funktionen $\mathrm{ps}_\nu^{-\mu}$ und $\mathrm{ps}_{\nu+p}^{-\mu}$ bzw. $\mathrm{ps}_{\nu+p}^{-\mu-1}$ erscheinen. Analog in (24), (25).

Die Bemerkungen b) und c) interessieren insbesondere für den Fall der Sphäroidfunktionen ganzer Ordnung und ganzen Grades. Wir verzichten

jedoch darauf, derartige Formeln explizit zu notieren. Ihre Gewinnung bietet mit den gegebenen Bemerkungen und den Formeln von [1] keine Schwierigkeiten.

d) Nach [1], 3.544. sind die erhaltenen Integralrelationen äquivalent zu Entwicklungen der Kerne $v_1(\xi,\eta)$, $v_2(\xi,\eta)$ nach Produkten von Sphäroidfunktionen

$$S_{\nu+p}^{\mu\,(j)}(\xi;\gamma)\,\tilde{\mathrm{Q}}\mathrm{s}_{\nu+p}^{\mu}(\eta;\gamma^2)$$

bzw.

$$S_{\nu+p}^{\mu+1\,(j)}(\xi;\gamma)\,\tilde{\mathrm{Q}}\mathrm{s}_{\nu+p}^{\mu+1}(\eta;\gamma^2)$$

mit ungeradem p. Im Falle b), c) kann man hier wieder die $\tilde{\mathrm{Q}}\mathrm{s}_{\nu+p}^{\mu}$, $\tilde{\mathrm{Q}}\mathrm{s}_{\nu+p}^{\mu+1}$ durch die $\mathrm{ps}_{\nu+p}^{-\mu}$, $\mathrm{ps}_{\nu+p}^{-\mu-1}$ ersetzen.

25. *Folgerungen*

Schreibt man die beiden Integralrelationen je für $j=3$ und $j=4$ auf, so kann man durch Auflösung linearer Gleichungssysteme Ausdrücke für die Integrale

(30) $$\oint \frac{\eta}{\xi^2-\eta^2}\,\tilde{\mathrm{Q}}\mathrm{s}_{\nu}^{\mu}(\eta;\gamma^2)\,\tilde{\mathrm{Q}}\mathrm{s}_{-\nu-p-1}^{\mu}(\eta;\gamma^2)\,d\eta,$$

(31) $$\oint \frac{1-\eta^2}{\xi^2-\eta^2}\,\tilde{\mathrm{Q}}\mathrm{s}_{\nu}^{\mu\prime}(\eta;\gamma^2)\,\tilde{\mathrm{Q}}\mathrm{s}_{-\nu-p-1}^{\mu}(\eta;\gamma^2)\,d\eta,$$

(32) $$\oint \frac{(\eta^2-1)^{\frac{1}{2}}}{\xi^2-\eta^2}\,\tilde{\mathrm{Q}}\mathrm{s}_{\nu}^{\mu}(\eta;\gamma^2)\,\tilde{\mathrm{Q}}\mathrm{s}_{-\nu-p-1}^{\mu+1}(\eta;\gamma^2)\,d\eta,$$

(33) $$\oint \frac{(\eta^2-1)^{\frac{1}{2}}\eta}{\xi^2-\eta^2}\,\tilde{\mathrm{Q}}\mathrm{s}_{\nu}^{\mu\prime}(\eta;\gamma^2)\,\tilde{\mathrm{Q}}\mathrm{s}_{-\nu-p-1}^{\mu+1}(\eta;\gamma^2)\,d\eta$$

gewinnen (vgl. [1], 2.83. für Mathieusche Funktionen). Man erhält für sie der Reihe nach

(34) $$\left\{\frac{\gamma^2}{2i}(-1)^{\frac{p+1}{2}}\oint \eta\,\tilde{\mathrm{Q}}\mathrm{s}_{\nu}^{\mu}(\eta;\gamma^2)\,\tilde{\mathrm{Q}}\mathrm{s}_{-\nu-p-1}^{\mu}(\eta;\gamma^2)\,d\eta\times\right.$$
$$\left.\times\left[S_{\nu+p}^{\mu\,(3)}(\xi;\gamma)\,S_{\nu}^{\mu\,(4)}(\xi;\gamma)-S_{\nu+p}^{\mu\,(4)}(\xi;\gamma)\,S_{\nu}^{\mu\,(3)}(\xi;\gamma)\right],\right.$$

(35) $$\left\{\frac{\gamma^2}{2i}(-1)^{\frac{p+1}{2}}\frac{(\xi^2-1)}{\xi}\oint \eta\,\tilde{\mathrm{Q}}\mathrm{s}_{\nu}^{\mu}(\eta;\gamma^2)\,\tilde{\mathrm{Q}}\mathrm{s}_{-\nu-p-1}^{\mu}(\eta;\gamma^2)\,d\eta\times\right.$$
$$\left.\times\left[S_{\nu+p}^{\mu\,(4)}(\xi;\gamma)\,S_{\nu}^{\mu\,(3)\prime}(\xi;\gamma)-S_{\nu+p}^{\mu\,(3)}(\xi;\gamma)\,S_{\nu}^{\mu\,(4)\prime}(\xi;\gamma)\right],\right.$$

(36) $$\left\{\frac{\gamma^2}{2i}(-1)^{\frac{p+1}{2}}\frac{(\xi^2-1)^{\frac{1}{2}}}{\xi}\oint (\eta^2-1)^{\frac{1}{2}}\,\tilde{\mathrm{Q}}\mathrm{s}_{\nu}^{\mu}(\eta;\gamma^2)\,\tilde{\mathrm{Q}}\mathrm{s}_{-\nu-p-1}^{\mu+1}(\xi;\gamma^2)\,d\eta\times\right.$$
$$\left.\times\left[S_{\nu+p}^{\mu+1\,(3)}(\xi;\gamma)\,S_{\nu}^{\mu\,(4)}(\xi;\gamma)-S_{\nu+p}^{\mu+1\,(4)}(\eta;\gamma)\,S_{\nu}^{\mu\,(3)}(\xi;\gamma)\right],\right.$$

(37) $$\left\{\frac{\gamma^2}{2i}(-1)^{\frac{p+1}{2}}(\xi^2-1)^{\frac{1}{2}}\oint (\eta^2-1)^{\frac{1}{2}}\,\tilde{\mathrm{Q}}\mathrm{s}_{\nu}^{\mu}(\eta;\gamma^2)\,\tilde{\mathrm{Q}}\mathrm{s}_{-\nu-p-1}^{\mu+1}(\xi;\gamma^2)\,d\eta\times\right.$$
$$\left.\times\left[S_{\nu+p}^{\mu+1\,(3)}(\xi;\gamma)\,S_{\nu}^{\mu\,(4)\prime}(\xi;\gamma)-S_{\nu+p}^{\mu+1\,(4)}(\xi;\gamma)\,S_{\nu}^{\mu\,(3)\prime}(\xi;\gamma)\right]-\right.$$
$$\left.-\frac{\mu}{\xi^2-1}\oint (\eta^2-1)^{\frac{1}{2}}\,\tilde{\mathrm{Q}}\mathrm{s}_{\nu}^{\mu}(\eta;\gamma^2)\,\tilde{\mathrm{Q}}\mathrm{s}_{-\nu-p-1}^{\mu+1}(\eta;\gamma^2)\,d\eta.\right.$$

Bemerkenswert ist ein Vergleich des asymptotischen Verhaltens von (31) und (35) bzw. von (33) und (37). Er liefert mit [1], 3.65., (49), (50), (52)

$$(38)\quad \begin{cases} \oint (\eta^2-1)\,\tilde{Q}s_\nu^{\mu'}(\eta;\gamma^2)\,\tilde{Q}s_{-\nu-p-1}^{\mu}(\eta;\gamma^2)\,d\eta \\ = \left[\tfrac{1}{2}\left(\lambda_{\nu+p}^{\mu}(\gamma^2)-\lambda_\nu^\mu(\gamma^2)\right)-1\right]\oint \eta\,\tilde{Q}s_\nu^{\mu}(\eta;\gamma^2)\,\tilde{Q}s_{-\nu-p-1}^{\mu}(\eta;\gamma^2)\,d\eta \end{cases}$$

und

$$(39)\quad \begin{cases} \oint (\eta^2-1)^{\frac{1}{2}}\,\eta\,\tilde{Q}s_\nu^{\mu'}(\eta;\gamma^2)\,\tilde{Q}s_{-\nu-p-1}^{\mu+1}(\eta;\gamma^2)\,d\eta \\ = \left[\tfrac{1}{2}\left(\lambda_{\nu+p}^{\mu+1}(\gamma^2)-\lambda_\nu^\mu(\gamma^2)\right)-1\right]\oint (\eta^2-1)^{\frac{1}{2}}\,\tilde{Q}s_\nu^{\mu}(\eta;\gamma^2)\,\tilde{Q}s_{-\nu-p-1}^{\mu+1}(\eta;\gamma^2)\,d\eta - \\ \qquad -\mu\oint (\eta^2-1)^{-\frac{1}{2}}\,\tilde{Q}s_\nu^{\mu}(\eta;\gamma^2)\,\tilde{Q}s_{-\nu-p-1}^{\mu+1}(\eta;\gamma^2)\,d\eta. \end{cases}$$

3. Analoges für andere Klassen spezieller Funktionen

31. Zylinderfunktionen

311. Die Methode von 1. kann natürlich auch auf die Funktionen[3])

$$\psi_\nu^{(j)}(r)=\left(\frac{\pi}{2r}\right)^{\frac{1}{2}}\mathfrak{Z}_{\nu+\frac{1}{2}}^{(j)}(r)$$

angewandt werden. Die entsprechende Formel erhält man jedoch sofort aus (9) mit $\gamma\to 0$ gemäß [1], 3.65., (54):

$$(40)\quad \left[\nu_1(\nu_1+1)-\nu_2(\nu_2+1)\right]\int_{\infty e^{i0}}^{(0+)}\psi_{\nu_1}^{(1)}(r)\,\psi_{\nu_2}^{(1)}(r)\,dr=(e^{2\pi i(\nu_1+\nu_2)}-1)\sin\frac{\nu_1-\nu_2}{2}\pi.$$

312. Die Methode von 2. hat schon WHITTAKER [5] auf Bessel-Funktionen angewandt.

Wir schließen einfach so:

Mit

$$u(r,\varphi)=\mathfrak{Z}_\nu^{(j)}(r)\,e^{i\nu\varphi}$$

ist auch

$$\begin{aligned}\frac{\partial u}{\partial x_1}+i\frac{\partial u}{\partial x_2}&=e^{i\varphi}\left(\frac{\partial u}{\partial r}+\frac{i}{r}\frac{\partial u}{\partial\varphi}\right)\\&=\left[\mathfrak{Z}_\nu^{(j)'}(r)-\frac{\nu}{r}\mathfrak{Z}_\nu^{(j)}(r)\right]e^{i(\nu+1)\varphi}\end{aligned}$$

Lösung der zweidimensionalen Schwingungsgleichung

$$\Delta u+u=0.$$

Daher steht in der eckigen Klammer eine Zylinderfunktion zum Index $\nu+1$. Vergleich des asymptotischen Verhaltens zeigt dann unmittelbar, daß dies

$$-\mathfrak{Z}_{\nu+1}^{(j)}(r)$$

ist.

Die Methode liefert also hier — in überaus einfacher Weise — die Rekursionsformeln der Zylinderfunktionen.

[3]) Bezeichnungen nach [1].

32. *Kugelfunktionen*

Hier interessiert nur die Methode von 2. Man kann natürlich oben den Grenzübergang $\gamma \to 0$ machen. Mit Hilfe von (38), (39) erhält man dann aus den Integralrelationen die Rekursionsformeln der Funktionen $\psi_\nu^{(j)}(r)$.

Interessanter ist ein Versuch zur Herleitung der Rekursionsformeln der Kugelfunktionen. Wir skizzieren nur den Weg. Er erweist sich zum Teil einfacher als manche der üblichen Herleitungen.

Mit

$$x_1 = r(1-\eta^2)^{\frac{1}{2}} \cos\varphi, \qquad x_2 = r(1-\eta^2)^{\frac{1}{2}} \sin\varphi, \qquad x_3 = r\eta$$

gilt

$$\frac{\partial u}{\partial x_3} = \eta \frac{\partial u}{\partial r} + \frac{1}{r}(1-\eta^2)\frac{\partial u}{\partial \eta},$$

$$\frac{\partial u}{\partial x_1} \pm i \frac{\partial u}{\partial x_2} = e^{\pm i\varphi}\left[(1-\eta^2)^{\frac{1}{2}}\left(\frac{\partial u}{\partial r} - \frac{\eta}{r}\frac{\partial u}{\partial \eta}\right) \pm i \frac{1}{r(1-\eta^2)^{\frac{1}{2}}}\frac{\partial u}{\partial \varphi}\right].$$

Wir wählen

$$u = r^\nu \mathfrak{P}_\nu^\mu(\eta)\, e^{i\mu\varphi}$$

als Lösung von $\Delta u = 0$. Dann sind

$$v_1(r,\eta) = (1-\eta^2)\mathfrak{P}_\nu^{\mu'}(\eta)\frac{1}{r}r^\nu + \eta\,\mathfrak{P}_\nu^\mu(\eta)\,\nu r^{\nu-1}$$

und

$$v_2(r,\eta) = -\eta(1-\eta^2)^{\frac{1}{2}}\mathfrak{P}_\nu^{\mu'}(\eta)\frac{1}{r}r^\nu + (1-\eta^2)^{\frac{1}{2}}\mathfrak{P}_\nu^\mu(\eta)\,\nu r^{\nu-1} \mp \mu(1-\eta^2)^{-\frac{1}{2}}\mathfrak{P}_\nu^\mu(\eta)\frac{1}{r}r^\nu$$

in r, η separiert.

Betrachten wir zunächst den ersten Kern, so wird nach dem Vorstehenden

$$(1-\eta^2)\mathfrak{P}_\nu^{\mu'}(\eta) + \nu\eta\,\mathfrak{P}_\nu^\mu(\eta)$$

eine Kugelfunktion zu den Indizes $\nu - 1$ und μ. Untersuchung des Verhaltens bei $\eta = +1$:

$$\mathfrak{P}_\nu^\mu(\eta) = \frac{1}{\Gamma(1-\mu)}\left(\frac{\eta-1}{\eta+1}\right)^{\frac{\mu}{2}}\left(1 + P(\eta-1)\right)$$

liefert dann sofort, daß dies

$$(\nu+\mu)\,\mathfrak{P}_{\nu-1}^\mu(\eta)$$

ist. So hat man die grundlegende Rekursion bezüglich ν:

$$(1-\eta^2)\mathfrak{P}_\nu^{\mu'}(\eta) + \nu\eta\,\mathfrak{P}_\nu^\mu(\eta) = (\nu+\mu)\,\mathfrak{P}_{\nu-1}^\mu(\eta).$$

Beachtet man $\mathfrak{P}_\nu^\mu = \mathfrak{P}_{-\nu-1}^\mu$, $\mathfrak{P}_{\nu-1}^\mu = \mathfrak{P}_{-\nu}^\mu$ und schreibt man dann statt $-\nu-1$ wieder ν, so kommt auch

$$(1-\eta^2)\mathfrak{P}_\nu^{\mu'}(\eta) - (\nu+1)\eta\,\mathfrak{P}_\nu^\mu(\eta) = -(\nu-\mu+1)\,\mathfrak{P}_{\nu+1}^\mu(\eta).$$

Nun zum zweiten Kern! Hier wird analog

$$(\eta^2-1)^{-\frac{1}{2}}\left[(\eta^2-1)\left(\eta\,\mathfrak{P}_\nu^{\mu'}(\eta) - \nu\,\mathfrak{P}_\nu^\mu(\eta)\right) \mp \mu\,\mathfrak{P}_\nu^\mu(\eta)\right]$$

eine Kugelfunktion zu den Indizes $\nu-1, \mu \pm 1$, und zwar, wie das Verhalten bei $\eta=+1$ zeigt, gerade

$$\mathfrak{P}_{\nu-1}^{\mu+1}(\eta)$$

bzw.

$$(\nu+\mu)(\nu+\mu-1)\,\mathfrak{P}_{\nu-1}^{\mu-1}(\eta).$$

Man erhält so auch mit entsprechenden Umformungen die Rekursionen bezüglich μ.

33. Mathieusche Funktionen

Die Methode von 2. ist in [*1*], 2.83. durchgeführt worden. Man kann überdies oben $\mu=\frac{1}{2}$ setzen und gemäß [*1*], 3.536. auf Mathieusche Funktionen umrechnen.

4. Zur Frage der Existenz von Rekursionsformeln

Wie wir in 31. und 32. gesehen haben, führt der Gedankengang von 2. für Zylinderfunktionen und Kugelfunktionen (im letzteren Falle hier anscheinend erstmals durchgeführt) in sehr einfacher und natürlicher Weise zu den Rekursionsformeln dieser Funktionen. Dagegen liefert die Methode für die Mathieuschen Funktionen und Sphäroidfunktionen nur gewisse Relationen, in denen Integrale über Produkte dieser Funktionen auftreten. Man erkennt sogleich, woran dies liegt: Die Kerne sind bei den Zylinder- und Kugelfunktionen in einfacher Weise ausgeartet, nicht dagegen bei den höheren Funktionen.

Es dürfte nun kaum sinnvoll, im Gegenteil offenbar irreführend sein, wenn man die entsprechenden Relationen für die Mathieuschen Funktionen und Sphäroidfunktionen ebenfalls als „Rekursionsformeln" bezeichnen wollte. Dies geschieht leider bei WHITTAKER [*5*] und MARX [*3*], [*4*].

Wohl erhält man Relationen der Form

$$a(\xi;\nu,\gamma^2,\mu^2)\,S_\nu^{\mu(j)'}(\xi;\gamma)+b(\xi;\nu,\gamma^2,\mu^2)\,S_\nu^{\mu(j)}(\xi;\gamma)=S_{\nu+1}^{\mu+1(j)}(\xi;\gamma).$$

Jedoch sind die Koeffizienten a, b durch die zu verknüpfenden Sphäroidfunktionen selbst ausgedrückt. Die von MARX [*3*], S. 269 geäußerte Hoffnung, für sie einfachere Ausdrücke zu finden, erweist sich durch die folgenden Bemerkungen als trügerisch.

Bezüglich γ^2 sind $\lambda_\nu^\mu(\gamma^2)$, $\lambda_{\nu+1}^\mu(\gamma^2)$ im großen unendlich vieldeutige algebroide Funktionen, und zwar verschiedene, damit zugleich auch $S_\nu^{\mu(j)}(\xi;\gamma)$, $S_{\nu+1}^{\mu(j)}(\xi;\gamma)$. Man muß annehmen — erste numerische Resultate für die Eigenwerte $a_m(h^2)$, $b_m(h^2)$ der Mathieuschen Differentialgleichung legen das nahe —, daß jede der beiden Funktionen im großen in der γ^2-Ebene Spurpunkte von Verzweigungsstellen besitzt, die nicht auch solche der anderen sind, und zwar unendlich viele, derart, daß man unabhängig durch geeignete Umlaufswege von ν zu jedem $\nu+2k$ und von $\nu+1$ zu jedem $\nu+2l+1$ übergehen kann (k, l ganz). Dann müssen aber die Funktionen a und b dieselben Eigenschaften haben, also ebenfalls alle diese Verzweigungen von $\lambda_\nu^\mu(\gamma^2)$ und $\lambda_{\nu+1}^\mu(\gamma^2)$ liefern. So besteht kaum die Aussicht, für a und b Ausdrücke zu finden, die von Sphäroidfunktionen bzw. von ihren Eigenwerten unabhängig wären.

Das ließe jedoch die Möglichkeit zu, daß a und b sich als einfache Funktionen von ξ^{-2} erwiesen und sich entsprechende Konstanten durch $\lambda_\nu^\mu(\gamma^2)$, $\lambda_{\nu+1}^\mu(\gamma^2)$ ausdrücken ließen. Jedoch erscheint auch dies bei näherer Betrachtung von (30) bis (37) zusammen mit [1], 3.65., (49), (50), (52) als recht unwahrscheinlich.

Analoges gilt für die „Rekursionsformeln" bezüglich μ gemäß (23*), (25), (29) und für „Rekursionsformeln" der $\lambda_\nu^\mu(\gamma^2)$, wie sie MARX [3] im Zusammenhang mit dem Spezialfall von (9) hinschreibt.

Literatur

[1] MEIXNER, J., u. F. W. SCHÄFKE: Mathieusche Funktionen und Sphäroidfunktionen. In Grundlehren der mathematischen Wissenschaften, Bd. 71. Berlin-Göttingen-Heidelberg: Springer 1954. — [2] FLAMMER, C.: Prolate spheroidal wave functions. Stanford Research Institute Technical Report No. 16, 1952. — [3] MARX, I.: Recurrence relations for prolate spheroidal wave functions. J. Math. Phys. **32**, 269—275 (1954). — [4] MARX, I.: On the structure of recurrence relations. Michigan math. J. **2**, 45—50 (1954). — [5] WHITTAKER, E. T.: On the recurrence formulae for Mathieu functions. J. London Math. Soc. **4**, 88—96 (1929).

Mainz, Mathematisches Institut der Universität

(Eingegangen am 27. November 1956)

SCHÄFKE, F. W.
Math. Zeitschr. Bd. 67, S. 252—266 (1957)

Bemerkungen über Variationsrechnung, implizite Funktionen und Lagrangesche Multiplikatoren

Von

FRIEDRICH WILHELM SCHÄFKE

1. Meine erste Bemerkung knüpft an die Darstellung der Grundlagen der Variationsrechnung in dem bekannten Werk von CARATHEODORY[1]) an. C. gelingt der Aufbau der Theorie der einfachsten Grundintegrale

$$\int_{t_0}^{t_1} L\big(t, x_1(t), \ldots, x_n(t), x_1'(t), \ldots, x_n'(t)\big)\, dt$$

für mehrere gesuchte Funktionen bei festen Enden weitgehend allein unter den in gewisser Beziehung natürlichen Voraussetzungen

(V_1) *$L, L_{x_i}, L_{x_i'}, L_{x_i x_k}, L_{x_i x_k'}, L_{x_i' x_k'}$ seien stetig in einem Gebiete $\mathfrak{G}$ des $(2n+1)$-dimensionalen Raumes der Variablen $t, x_1, \ldots, x_n, x_1', \ldots, x_n'$.*

Er sieht sich jedoch zu schärferen Voraussetzungen genötigt und nimmt hier L nach allen $2n+1$ Variablen zweimal stetig differenzierbar an, sowie er zur Einführung einer Hamilton-Funktion, den kanonischen Gleichungen und zur Hamilton-Jacobischen Theorie kommt.

Überlegt man sich den Grund, so erkennt man, daß er allein in dem Satz über implizite Funktionen bzw. Umkehrung eines Funktionensystems liegt. Er wird angewandt, um statt der x_i' die

$$y_i = L_{x'}(t, x_1, \ldots, x_n, x_1', \ldots, x_n') \tag{1}$$

einzuführen, wozu man eben die Umkehrung

$$x_i' = \varphi_i(t, x_1, \ldots, x_n, y_1, \ldots, y_n) \tag{2}$$

braucht. C. wendet hier den klassischen Satz an: Sind die $L_{x_i'}$ nach allen Variablen stetig differenzierbar und ist die Determinante der $L_{x_i' x_k'}$ nicht 0, so ist „im Kleinen" eindeutige Umkehrung durch Funktionen φ_i möglich, die ebenfalls nach allen Variablen stetig differenzierbar sind. Es kann dann

$$H(t, x_i, y_i) = -L(t, x_i, \varphi_i) + \sum_j y_j \varphi_j$$

gebildet werden und erweist sich als zweimal stetig differenzierbar bezüglich aller Variablen.

[1]) CARATHEODORY, C.: Variationsrechnung, insbes. S. 190ff. Leipzig u. Berlin 1935.

Hier erhebt sich nun die Frage, ob es nicht gelingt, die Theorie konsequent und einheitlich unter den genannten „natürlichen" Voraussetzungen weiterzuführen, um so auch an den erwähnten Stellen etwa die Theorie der akzessorischen Variationsprobleme — wie es C. weitgehend anstrebt — einzubeziehen, ohne Modifikationen nötig zu haben. Wie man sieht, kommt dies auf die Frage nach einem entsprechend in Voraussetzung und Behauptung abgeschwächten Umkehrsatz für Funktionensysteme hinaus.

Daß dies möglich ist, legt schon die Betrachtung des einfachsten Falles, nämlich der Auflösung von

$$f(t, x) = 0$$

nahe. Hier ist die Existenz einer eindeutigen stetigen auflösenden Funktion $x(t)$ um (t_0, x_0) schon gesichert, wenn nur f und f_x in der Umgebung stetig sind und

$$f(t_0, x_0) = 0, \quad f_x(t_0, x_0) \neq 0$$

gilt. Das lehrt die übliche Beweismethode (Monotonie von f bezüglich x).

Der angestrebte Satz soll im folgenden sogleich in einer Form gegeben werden, die in zwei Richtungen allgemeiner ist: einmal werden Abbildungen zwischen Banach-Räumen betrachtet, zum anderen werden etwa die Sätze über die Einbettbarkeit von Extremalenbögen in Felder mit umfaßt[2]).

Zur Formulierung wird die von V. STOLTZ herrührende und von FRÉCHET in die Funktionalanalysis übertragene Definition der Differentiation verwandt:

$\mathfrak{L}_1, \mathfrak{L}_2$ seien reelle lineare normierte Räume, $\mathfrak{M}$ eine offene Menge in $\mathfrak{L}_1$, F eine Abbildung von $\mathfrak{M}$ in $\mathfrak{L}_2$. Es sei Λ_1 die Gesamtheit der linearen stetigen Abbildungen von $\mathfrak{L}_1$ in $\mathfrak{L}_2$, die in gewohnter Weise wieder als linearer normierter Raum aufgefaßt werde. — In $x_0 \in \mathfrak{M}$ heißt nun F differenzierbar, wenn ein $F'(x_0) \in \Lambda_1$ existiert (dann eindeutig), derart daß

$$F(x_0 + h) - F(x_0) = F'(x_0)\, h + o(\|h\|) \qquad (h \to 0)$$

gilt. — Ist F an jeder Stelle von $\mathfrak{M}$ differenzierbar, so ist F' eine Abbildung von $\mathfrak{M}$ in Λ_1 und man kann gegebenenfalls die zweite Abbildung betrachten. Es übertragen sich in sehr naheliegender Weise Kettenregel, Mittelwertsatz und Taylorscher Satz[3]).

Der Mittelwertsatz erscheint in der Form: Ist F eine differenzierbare Abbildung der offenen konvexen Menge $\mathfrak{M}$ von $\mathfrak{L}_1$ in $\mathfrak{L}_2$, f^* eine lineare stetige Abbildung von $\mathfrak{L}_2$ in P, den Banach-Raum der reellen Zahlen (f^* also Element des dualen Raumes $\mathfrak{L}_2^*$), dann gilt für $x_1 \neq x_2$ aus $\mathfrak{M}$

$$f^*(F(x_2) - F(x_1)) = f^* F'(x_{12})\,(x_2 - x_1)$$

mit einem

$$x_{12} = x_1 + \Theta(x_2 - x_1), \qquad 0 < \Theta < 1.$$

---

[2]) Vgl. z. B. C. BÓLZA: Vorlesungen über Variationsrechnung, S. 160ff. Leipzig u. Berlin 1909.

[3]) Vgl. z. B. LJUSTERNIK-SOBOLEV: Elemente der Funktionalanalysis (deutsche Übersetzung); insbes. Kap. VI. Berlin 1955.

Zum Beweise wird die reelle Funktion von τ $(0 \leq \tau \leq 1)$

$$f^* F(x_1 + \tau(x_2 - x_1))$$

betrachtet und die Kettenregel verwandt. Entsprechend wird für den Taylorschen Satz verfahren.

Der angekündigte umfassende Umkehrsatz ist der

Satz 1. *$\mathfrak{B}_0$, $\mathfrak{B}_1$, $\mathfrak{B}_2$ seien reelle Banach-Räume, die Variablen in ihnen mit t, x, y bezeichnet. Im kartesischen Produkt $\mathfrak{B}_0 \times \mathfrak{B}_1$ [4]) sei $\mathfrak{M}$ eine offene Menge. Sie werde durch die Abbildung F in $\mathfrak{B}_2$ abgebildet. $\mathfrak{C}$ sei eine kompakte Teilmenge von $\mathfrak{M}$. F sei in $\mathfrak{M}$ nach x partiell differenzierbar. F und die partielle Ableitung F_x seien in den Punkten von $\mathfrak{C}$ (also nicht nur „auf" $\mathfrak{C}$) stetig. Für jedes $(t_0, x_0) \in \mathfrak{C}$ möge $F_x(t_0, x_0)$ eine umkehrbar eindeutige Abbildung [5]) von $\mathfrak{B}_1$ auf $\mathfrak{B}_2$ vermitteln. Schließlich gelte für $(t_0, x_1) \in \mathfrak{C}$, $(t_0, x_2) \in \mathfrak{C}$ mit $x_1 \neq x_2$ stets $F(t_0, x_1) \neq F(t_0, x_2)$.*

Dann (1) *gibt es ein $r > 0$, derart daß für jedes Paar $(t_1, x_1) \neq (t_2, x_2)$ verschiedener Elemente von $\mathfrak{M}$ mit Abständen von $\mathfrak{C}$*

$$\delta\big((t_i, x_i), \mathfrak{C}\big) < r \qquad (i = 1, 2)$$

auch gilt

$$\big(t_1, F(t_1, x_1)\big) \neq \big(t_2, F(t_2, x_2)\big).$$

(2) *$\mathfrak{C}$ wird umkehrbar eindeutig abgebildet auf die kompakte Menge*

$$\tilde{\mathfrak{C}} = \{(t, y) \mid y = F(t, x),\ (t, x) \in \mathfrak{C}\}$$

aus $\mathfrak{B}_0 \times \mathfrak{B}_2$. (3) *Es gibt ein $\varrho > 0$, derart daß für jedes (t, y) mit*

$$(*) \qquad \delta\big((t, y), \tilde{\mathfrak{C}}\big) < \varrho$$

(eindeutig) ein x mit

$$\delta\big((t, x), \mathfrak{C}\big) < r, \qquad y = F(t, x)$$

existiert. (4) *Setzt man dies*

$$x = \Phi(t, y),$$

so ist Φ im Definitionsbereich (*) nach y partiell differenzierbar, und zwar

$$\Phi_y(t, y) = F_x(t, x)^{-1}.$$

(5) *Φ und Φ_y sind in den Punkten von $\tilde{\mathfrak{C}}$ stetig.*

Der Beweis knüpft an die von HILDEBRANDT und GRAVES zum Beweise ihres Satzes über implizite Funktionen verwandten Methoden [6]) an und soll in Abschnitt 2 gegeben werden.

---

[4]) Wir wählen $\|(t, x)\| = \max(\|t\|, \|x\|)$.

[5]) Dann ist nach einem bekannten Satz von S. BANACH auch $F_x(t_0, x_0)^{-1}$ (linear und) stetig.

[6]) HILDEBRANDT, T. H., and I. M. GRAVES: Implizit functions and their differentials in general analysis. Trans. Amer. Math. Soc. **29**, 127—153 (1927). Vgl. auch die Darstellung (ohne Zitat) in LUSTERNIK-SOBOLEV, l. c. Fußnote 3, S. 222ff. Die dort ge-

Hier sollen nur noch die Folgerungen des Satzes für die Begründung der Variationsrechnung, von der wir oben ausgingen, kurz genannt werden:

Unter den Voraussetzungen (V_1) folgt bei Nichtverschwinden der Determinante der $L_{x_i' x_k'}$ „im Kleinen" die Umkehrbarkeit von (1) durch Funktionen (2), die zusammen mit ihren partiellen Ableitungen erster Ordnung nach den x_i, y_i in allen $2n+1$ Variablen t, x_j, y_j stetig sind. [Man wird dabei zur Anwendung von Satz 1 in gewohnter Weise zu (1) die n Gleichungen $z_i = x_i$ hinzunehmen.]

Bildet man dann (3), so folgt natürlich

$$H_{x_j} = -L_{x_j}(t, x_i, \varphi_i), \qquad H_{y_j} = \varphi_j,$$

und mit (V_1) und den Eigenschaften der φ_j wiederum, daß

$$H, H_{x_i}, H_{y_i}, H_{x_i x_k}, H_{x_i y_k}, H_{y_i y_k}$$

bezüglich aller $2n+1$ Variablen t, x_j, y_j stetig sind. Das genügt völlig zum Aufbau der Hamilton-Jacobischen Theorie[7]).

Eine zweite Bemerkung schließt sich unmittelbar an.

Wie man sieht, ist mit Hilfe derartiger Sätze der Nachweis der Existenz einer Hamilton-Funktion stets nur „im Kleinen" möglich. Das ist recht unbefriedigend, da man ja vor allem Probleme „im Großen" im Auge hat. Die mir bekannten Darstellungen gehen entweder ganz darüber hinweg, oder sie postulieren die Existenz *einer* Hamilton-Funktion. Das entspricht wohl der klassischen Variations*rechnung*, die vor allem an Beispielen orientiert ist und bei der die Theorie oft nur allgemeine Formulierung des Vorgehens an Beispielen ist. Im Rahmen einer modernen strengen Variationstheorie jedoch, wie man sie etwa seit CARATHEODORY anstrebt, stellt die genannte Tatsache wohl eine gewisse Lücke dar.

Es zeigt sich nun, daß eine sehr befriedigende und überaus einfache Lösung dieser Aufgabe auf Grund eines entsprechenden Umkehrsatzes im Großen möglich ist: Hat man ein positiv-reguläres Variationsproblem, d.h. fordert man zusätzlich zu (V_1) noch

(V_2) *$\mathfrak{S}$ ist bezüglich der Variablen x_i' konvex. An jeder Stelle von $\mathfrak{S}$ ist die quadratische Form*

$$\sum_{i,k} L_{x_i' x_k'}(t, x_j, x_j')\, \xi_i \xi_k$$

positiv definit.

dann gilt:

Durch die Gln. (1) wird das Gebiet $\mathfrak{S}$ eineindeutig auf ein Gebiet $\mathfrak{T}$ des $(2n+1)$-dimensionalen Raumes der Variablen t, x_i, y_i abgebildet.

gebenen Sätze beziehen sich auf Punktausartung von $\mathfrak{C}$, liefern aber auch dann nicht den schärferen Inhalt unseres Satzes, wohl aber mit einer geläufigen Umformung umgekehrt. Beim Satze von LUSTERNIK-SOBOLEV, S. 222 muß offenbar die Voraussetzung der Stetigkeit von $\varphi(x, y)$ ergänzt werden. Ein Teilresultat in unserer Richtung in bezug auf Funktionensysteme wird bei J. HADAMARD, Lecons du calcul des variations, I, Paris 1910, S. 497, Anmerkung 2, erwähnt.

[7]) Vgl. dazu auch die Bemerkung 2 von C. CARATHEODORY, l. c. S. 51.

Dann hat man also „im Großen" *eine* Hamilton-Funktion.

Zugrunde liegt diesem Resultat ein einfacher Satz über Umkehrbarkeit im Großen, der als Verallgemeinerung der Tatsache erscheint, daß eine in einem Intervall differenzierbare Funktion mit überall positiver Ableitung eine entsprechende Umkehrfunktion besitzt. Es muß nur statt „Intervall" „konvexes Gebiet" und statt der Positivität eine Positiv-Definitheit gefordert werden. Natürlich läßt sich der Satz am bequemsten sofort für Banach-Räume formulieren:

Satz 2. *Sei $\mathfrak{B}$ ein reeller Banach-Raum, $\mathfrak{B}^*$ der duale Raum, F eine differenzierbare Abbildung eines konvexen Gebietes $\mathfrak{G}$ von $\mathfrak{B}$ in $\mathfrak{B}^*$. Diese habe die Eigenschaft, daß für alle x aus $\mathfrak{G}$ und alle $h \neq 0$ aus $\mathfrak{B}$ stets*

$$(F'(x)\,h)\,h \neq 0$$

gelte.

Dann gilt für $x_1 \neq x_2$ aus $\mathfrak{G}$ immer $F(x_1) \neq F(x_2)$.

Man beachte dazu, daß $F'(x)$ eine lineare stetige Abbildung von $\mathfrak{B}$ in $\mathfrak{B}^*$, daher $F'(x)h$ aus $\mathfrak{B}^*$ und somit $(F'(x)h)h$ eine reelle Zahl ist.

Der Beweis beruht auf dem oben genannten Mittelwertsatz:

$$\big(F(x_2) - F(x_1)\big)\,(x_2 - x_1) = \big(F'(x_{12})\,(x_2 - x_1)\big)\,(x_2 - x_1) \neq 0.$$

Es ist wohl unmittelbar ersichtlich, wie dieser Satz zu dem vorher notierten Ergebnis für die Variationstheorie führt.

Eine dritte und letzte Bemerkung betrifft die Frage nach der Möglichkeit, die Variationstheorie weitgehend in eine allgemeine Analysis einzubetten, die zugleich Extremalaufgaben für Zahlenfunktionen umfaßt. Lange bekannt — nämlich einer der Ausgangspunkte für den Begriff des Fréchetschen Differentials — ist natürlich folgendes:

Bei den o.B.d.A. wählbaren Enden

$$x_i(t_0) = x_i(t_1) = 0 \qquad (i = 1, 2, \ldots, n)$$

kann das Grundintegral

$$\int_{t_0}^{t_1} L\big(t, x_j(t), x_j'(t)\big)\,dt = F(x)$$

als Abbildung einer offenen Menge des Banach-Raumes der an den Enden verschwindenden in $[t_0, t_1]$ stetig differenzierbaren reellen Vektorfunktionen mit

$$\|x\| = \max_{t,j,k} \big(|x_j(t)|, |x_k'(t)|\big)$$

in den Banach-Raum P der reellen Zahlen aufgefaßt werden. Unter den Voraussetzungen (V_1)[8]) ist dann das „erste Differential"

$$F'(x)\,h = \int_{t_0}^{t_1} \sum_i \big(L_{x_i}(t, x_j(t), x_j'(t))\,h_i(t) + L_{x'}(t, x_j(t), x_j'(t))\,h_i'(t)\big)\,dt$$

[8]) Für die erste Variation benötigt man natürlich die zweiten Ableitungen nicht.

und nichts anderes als die „erste Variation", ebenso das „zweite Differential" (vgl. den entsprechenden Taylorschen Satz)

$$(F''(x)\,h)\,h = \int_{t_0}^{t_1} \sum_{i,k} (L_{x_i x_k} h_i h_k + 2 L_{x_i x_k'} h_i h_k' + L_{x_i' x_k'} h_i' h_k')\,dt$$

und nichts anderes als die „zweite Variation".

So folgt mancherlei, was oft für die Variationsrechnung etwas umständlich begründet wird, einfach aus den sehr naheliegenden Sätzen:

a) Hat eine differenzierbare Abbildung F einer offenen Menge $\mathfrak{M}$ eines reellen linearen normierten Raumes in den Raum P im Punkte x_0 ein relatives Minimum, so ist $F'(x_0) = 0$, das Nullfunktional.

b) Ist dabei F zweimal differenzierbar, so hat man zusätzlich für alle h des Raumes $(F''(x_0)\,h)\,h \geq 0$.

Mit Hilfe des Taylorschen Satzes kann man natürlich auch zu gewissen hinreichenden Bedingungen kommen. Sie erweisen sich jedoch in dieser Form als für die Variationstheorie kaum besonders handlich.

Interessanter und anscheinend im hier gegebenen Umfange noch nicht gelöst ist nun die Frage nach gemeinsamer Behandlung der Extremalprobleme mit Nebenbedingungen, also nach einem umfassenden Satz über Lagrangesche Multiplikatoren. Er müßte sowohl die Zahlenmultiplikatoren bei den isoperimetrischen Problemen (was allein kaum Schwierigkeiten macht) als auch etwa die Multiplikatorfunktionen bei holonomen Bindungen liefern[9]).

Dies erfüllt nun der folgende

Satz 3. *$\mathfrak{B}$ und $\mathfrak{B}_3$ seien reelle Banach-Räume. Die offene Menge $\mathfrak{M}$ von $\mathfrak{B}$ werde durch die stetig differenzierbaren Abbildungen f bzw. G in P (reelle Zahlen) bzw. $\mathfrak{B}_3$ abgebildet. Für $a \in \mathfrak{M}$ sei $G(a) = 0$. $f(x)$ habe in a ein relatives Minimum unter der Nebenbedingung $G(x) = 0$. $\mathfrak{B}$ lasse eine Zerlegung als nichttriviale direkte Summe $\mathfrak{B} = \mathfrak{B}_1 + \mathfrak{B}_2$ zweier Unterräume $\mathfrak{B}_1$, $\mathfrak{B}_2$ zu. Durch $G'(a)$ werde $\mathfrak{B}_1$ eineindeutig auf $\mathfrak{B}_3$ abgebildet. (Die dann stetige lineare Umkehrung, die $\mathfrak{B}_3$ auf $\mathfrak{B}_1$ abbildet, sei mit $G'(a)^{-1}$ bezeichnet.)*

Dann existiert ein λ aus dem dualen Raum $\mathfrak{B}_3^*$, derart daß

$$f'(a) = \lambda G'(a).$$

Es ist[10])

$$\lambda = f'(a)\,G'(a)^{-1}.$$

Sind zusätzlich f und G zweimal stetig differenzierbar, so hat man noch: Für jedes h aus $\mathfrak{B}$ mit $G'(a)\,h = 0$ gilt

$$[(f''(a) - \lambda G''(a))\,h]\,h \geq 0.$$

[9]) Unser Satz dürfte eine gewisse Ergänzung geben zu den Bestrebungen bei der Behandlung der Lagrangeschen Multiplikatoren, wie sie insbesondere H. Hahn, Monatsh. f. Math. **14**, 325—342 (1903); **17**, 295—304 (1906), verfolgt hat.

[10]) Das ist — vgl. die Definition von $G'(a)^{-1}$ — keine triviale Umformung.

Zur Erläuterung sollte gesagt werden: $\mathfrak{B} = \mathfrak{B}_1 + \mathfrak{B}_2$ bedeutet, daß jedes $x \in \mathfrak{B}$ eindeutig eine Darstellung $x = x_1 + x_2$ mit $x_1 \in \mathfrak{B}_1$, $x_2 \in \mathfrak{B}_2$ besitzt, wobei die x_i stetig (und linear) von x abhängen. Daraus folgt die Isomorphie $\mathfrak{B} \cong \mathfrak{B}_1 \times \mathfrak{B}_2$ im Sinne einer Isomorphie der linearen Räume und einer Äquivalenz der Normen (jede Norm relativ zur anderen beschränkt).

Wir führen den Beweis von Satz 3 unten in Abschnitt 3. Im folgenden deuten wir kurz mögliche Anwendungen an. Naheliegend ist natürlich die Spezialisierung auf Zahlfunktionen, auf die wir daher verzichten wollen.

Betrachten wir zuerst die Anwendung auf isoperimetrische Probleme. Neben L seien die Grundfunktionen $L_1, L_2, \ldots, L_k$ gegeben. Für alle diese sei (V_1) erfüllt. Wir betrachten als feste Enden o.B.d.A.

$$x_i(t_0) = x_i(t_1) = 0.$$

$\mathfrak{B}$ sei nun der oben genannte Banach-Raum der stetig differenzierbaren Vektorfunktionen mit dieser Eigenschaft. Wir setzen

$$f(x) = \int_{t_0}^{t_1} L\big(t, x_i(t), x_i'(t)\big)\, dt,$$

$$g_\varrho(x) = \int_{t_0}^{t_1} L_\varrho\big(t, x_i(t), x_i'(t)\big)\, dt,$$

$$G(x) = \big(g_1(x), \ldots, g_k(x)\big)$$

und wählen $\mathfrak{B}_3 = \mathsf{P}^{\times k}$, d.h. als k-dimensionalen Vektorraum. In einer geeigneten offenen Menge $\mathfrak{M}$ von $\mathfrak{B}$ (von $\mathfrak{S}$ abhängig) sind dann f und G zweimal stetig differenzierbar.

Es möge nun für $a \in \mathfrak{M}$ $f(x)$ ein relatives Minimum unter der Nebenbedingung $G(x) = 0$ besitzen. Dann unterscheiden wir zwei Fälle:

1) $G'(a)h$ $(h \in \mathfrak{B})$ erfüllt nicht den ganzen Raum $\mathfrak{B}_3$. Dann ist die Wertemenge ein echter Unterraum $\mathfrak{U}$. Im dualen Raume $\mathfrak{B}_3^* \cong \mathsf{P}^{\times k}$ gibt es ein $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_k) \neq 0$, derart daß für alle $u = (u_1, u_2, \ldots, u_k)$ aus $\mathfrak{U}$ gilt $\lambda u = \sum_1^k \lambda_i u_i = 0$. Dann gilt mit

$$(\lambda_0, \lambda_1, \ldots, \lambda_k) = (0, \lambda_1, \ldots, \lambda_k) \neq 0$$

die isoperimetrische Bedingung

$$\lambda_0 f'(a) + \lambda_1 g_1'(a) + \cdots + \lambda_k g_k'(a) = 0$$

(Nullfunktional).

2) $G'(a)h$ $(h \in \mathfrak{B})$ erfüllt ganz $\mathfrak{B}_3$. Dann wähle man $h_1, h_2, \ldots, h_k$ aus $\mathfrak{B}$ mit $g_i'(a)h_j = \delta_{ij}$. Die $h_1, \ldots, h_k$ sind linear unabhängig. Man wähle $\mathfrak{B}_1$ als den von diesen aufgespannten Unterraum und findet in bekannter Weise eine Zerlegung $\mathfrak{B} = \mathfrak{B}_1 + \mathfrak{B}_2$. Damit sind dann die Voraussetzungen von Satz 3 gegeben. Man hat somit ein

$$\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_k) \neq 0$$

aus $\mathfrak{B}_3^*$ mit $f'(a) = \lambda G'(a)$, d.h.

$$f'(a) = \lambda_1 g_1'(a) + \cdots + \lambda_k g_k'(a).$$

Das ist wieder die isoperimetrische Bedingung. Man kann sie auf die ersten Variationen umschreiben und hat sie dann in Form der Eulerschen Differentialgleichungen. Entsprechend kann man nun sofort die Bedingung zweiter Ordnung von Satz 3 für die zweiten Variationen aufschreiben.

Diskutieren wir weiter die Anwendung von Satz 3 auf Variationsprobleme mit holonomen Bindungen. L, $\mathfrak{S}$, $\mathfrak{B}$, $\mathfrak{M}$, f mögen dieselbe Bedeutung haben und dieselben Voraussetzungen erfüllen wie eben. Es sei $n \geq 2$. Zusätzlich seien die $L_{x_i'}$ in $\mathfrak{S}$ vorhanden und stetig. Dann wählen wir $\mathfrak{B}_3$ als den Banach-Raum der stetig differenzierbaren Funktionen $y(t)$ in $[t_0, t_1]$ mit $y_0(t) = y(t_1) = 0$ und

$$\|y\| = \max(|y(t)|, |y'(t)|).$$

In der Projektion von $\mathfrak{S}$ auf den $(n+1)$-dimensionalen Raum der Variablen t, $x_1, \ldots, x_n$ sei die Funktion $M(t, x_1, x_2, \ldots, x_n)$ mit den Ableitungen

$$M_t, M_{x_i}, M_{x_i t}, M_{x_i x_k}, M_{x_i x_k t}, M_{x_i x_k x_l}$$

stetig und

$$M(t_0, 0, \ldots, 0) = M(t_1, 0, \ldots, 0) = 0.$$

Durch

$$y = G(x) = M\big(t, x_1(t), \ldots, x_n(t)\big)$$

wird dann eine zweimal stetig differenzierbare Abbildung von $\mathfrak{M}$ in $\mathfrak{B}_3$ vermittelt.

In a aus $\mathfrak{M}$ möge nun f ein relatives Minimum bei der Nebenbedingung $G(x) = 0$ haben[11]). Es gebe einen Index p aus $1, 2, \ldots, k$ derart, daß

$$M_{x_p}\big(t, a_1(t), \ldots, a_n(t)\big) \neq 0 \qquad (t_0 \leq t \leq t_1)$$

sei. $a(t)$ sei zweimal stetig differenzierbar.

Dann kann man $\mathfrak{B}_1$ als den Unterraum von $\mathfrak{B}$ wählen, in dem alle $x_i(t) \equiv 0$ für $i \neq p$ sind und $\mathfrak{B}_2$ entsprechend als den mit $x_p(t) \equiv 0$. Für $h \in \mathfrak{B}_1$ wird dann

$$G'(a)\,h = M_{x_p}\big(t, a_1(t), \ldots, a_n(t)\big)\, h_p(t).$$

Damit hat man eine eineindeutige Abbildung zwischen $\mathfrak{B}_1$ und $\mathfrak{B}_3$, und man kann Satz 3 anwenden.

Dazu ist es nun wesentlich, die Struktur des Funktionals λ zu kennen. Sie wird einfach aus Satz 3 entnommen. Man beachtet, daß Anwendung von $G'(a)^{-1}$ auf $y \in \mathfrak{B}_3$ eine Multiplikation mit einer stetig differenzierbaren Funktion bedeutet, ferner, daß Anwendung von $f'(a)$ auf $x \in \mathfrak{B}_1$ die Multiplikation von x_p mit einer stetigen Funktion, nämlich mit

$$L_{x_p} - (L_{x_p'})'\big|_{x=a}$$

[11]) Es ist also speziell $G(a) = 0$, d.h. $M(t, a_1(t), \ldots, a_n(t)) \equiv 0$.

[wozu man die Stetigkeit von $L_{x_p' t}$ und $a''(t)$ beachtet], und nachfolgende Integration über $[t_0, t_1]$ bedeutet. Damit ist das Funktional λ nichts anderes als die Multiplikation mit einer stetigen Funktion $\lambda(t)$ und nachfolgende Integration.

Man kommt so in der gewohnten Weise auf die Eulerschen Differentialgleichungen

$$(N_{x_i'})' = N_{x_i}$$

mit

$$N = L(t, x_1, \ldots, x_n') - \lambda(t) M(t, x_1, \ldots, x_n)$$

für $x = a$. Weiter kann man nun sofort die Bedingung zweiter Ordnung hinschreiben.

Diese Andeutungen für die Anwendungen von Satz 3 mögen genügen. Man sieht schnell, wie man entsprechend für mehrere holonome Bindungen und schließlich auch für nicht-holonome Bindungen vorgehen kann.

2. Wir führen im folgenden den Beweis von Satz 1. Wir schicken, um später den Gedankengang weniger unterbrechen zu müssen, zwei Hilfssätze voraus.

Hilfssatz 1. *A, B seien stetige lineare Abbildungen eines Banach-Raumes $\mathfrak{B}_1$ in den Banach-Raum $\mathfrak{B}_2$. Es existiere A^{-1}* [12]*. Man habe*

$$\|A^{-1}\| \leq \frac{1}{a}, \qquad \|B - A\| \leq \alpha < a.$$

Dann existiert auch B^{-1} und man hat

$$\|B^{-1}\| \leq \frac{1}{a - \alpha}, \qquad \|B^{-1} - A^{-1}\| \leq \frac{\alpha}{a(a - \alpha)}.$$

Beweis. Man schreibt

$$B = \left(E + (B - A) A^{-1}\right) A,$$

wo E der Einheitsoperator in $\mathfrak{B}_2$ und $(B - A) A^{-1}$ nach Voraussetzung stetige lineare Abbildung von $\mathfrak{B}_2$ in sich mit

$$\|(B - A) A^{-1}\| \leq \frac{\alpha}{a} < 1$$

ist. Dann existiert aber bekanntlich

$$\left(E + (B - A) A^{-1}\right)^{-1} = E - [(B - A) A^{-1}] + [(B - A) A^{-1}]^2 - + \cdots$$

mit

$$\|(E + (B - A) A^{-1})^{-1}\| \leq \frac{1}{1 - \frac{\alpha}{a}},$$

$$\|(E + (B - A) A^{-1})^{-1} - E\| \leq \frac{\frac{\alpha}{a}}{1 - \frac{\alpha}{a}}.$$

[12] Das heißt A bilde $\mathfrak{B}_1$ und $\mathfrak{B}_2$ eineindeutig aufeinander ab, so daß nach S. BANACH auch A^{-1} linear und stetig ist.

Daraus folgen nun sofort die Existenz von

$$B^{-1} = A^{-1}\left(E + (B - A)\,A^{-1}\right)^{-1}$$

und die behaupteten Abschätzungen.

Hilfssatz 2. Die differenzierbare Abbildung F bilde die Kugel $\|x - a\| < r_1$ des Banach-Raumes $\mathfrak{B}_1$ in den Banach-Raum $\mathfrak{B}_2$ ab. Es sei $F(a) = b$. A sei eine stetige lineare Abbildung von $\mathfrak{B}_1$ in $\mathfrak{B}_2$. Es existiere A^{-1} [12]) mit $\|A^{-1}\| \leq \frac{1}{m}$. Für alle x mit $\|x - a\| < r_1$ gelte

$$\|F'(x) - A\| \leq \frac{m}{2}.$$

Dann nimmt $F(x)$ in $\|x - a\| < r_1$ keinen Wert zweimal und jeden Wert y mit $\|y - b\| < \frac{m}{2} r_1$ einmal an. Man hat für x_i aus $\|x - a\| < r_1$

$$\frac{m}{2}\|x_2 - x_1\| \leq \|F(x_2) - F(x_1)\|.$$

Beweis. Wir nehmen $a = 0$, $b = 0$ an, was wegen des möglichen Übergangs zu $F(x + a) - b$ an Stelle von $F(x)$ keine Einschränkung darstellt. Wir setzen

$$(\circ) \qquad F(x) - A\,x = \varphi(x).$$

Dann gilt für x_1, x_2 mit $\|x_i\| < r_1$ und ein $f^* \in \mathfrak{B}_2^*$ nach dem Mittelwertsatz mit einem $x_{12} = x_1 + \vartheta(x_2 - x_1)$, $0 < \vartheta < 1$,

$$f^*\left(\varphi(x_2) - \varphi(x_1)\right) = f^*\left(F'(x_{12}) - A\right)(x_2 - x_1).$$

Wir wählen speziell (Hahn-Banach)

$$\|f^*\| = 1, \quad f^*(\varphi(x_2) - \varphi(x_1)) = \|\varphi(x_2) - \varphi(x_1)\|$$

und erhalten

$$(*) \qquad \|\varphi(x_2) - \varphi(x_1)\| \leq \frac{m}{2}\|x_2 - x_1\|.$$

Daraus folgt mit $(\circ)$ und $\|A^{-1}\| \leq \frac{1}{m}$, also $\|A(x_2 - x_1)\| \geq m\|x_2 - x_1\|$, sofort die letzte und so erst recht die erste Behauptung des Satzes. Wir bilden weiter

$$\psi(x) = -A^{-1}\varphi(x),$$

für das nun mit $\|x_i\| < r_1$ gilt

$$(**) \qquad \|\psi(x_2) - \psi(x_1)\| \leq \frac{1}{2}\|x_2 - x_1\|.$$

Jetzt ist offenbar $y = F(x)$ äquivalent mit

$$x = A^{-1}y + \psi(x).$$

Daraus wird die zweite Behauptung erhalten, indem man das Iterationsverfahren anwendet:

$$x_0 = 0, \qquad x_{n+1} = A^{-1}y + \psi(x_n) \qquad (n = 0, 1, 2, \ldots).$$

Man hat für $\|y\| < \frac{m}{2} r_1$

$$\|x_1 - x_0\| < \frac{1}{2} r_1$$

und

$$\|x_{n+1} - x_n\| \leq \frac{1}{2} \|x_n - x_{n-1}\| \qquad (n = 1, 2, 3, \ldots).$$

Also hat man $\|x_n\| < r_1$ $(n = 0, 1, 2, \ldots)$, die x_n konvergieren gegen ein x mit $\|x\| < r_1$ und es gilt $x = A^{-1}y + \psi(x)$, d.h. $y = F(x)$.

Beweis von Satz 1. *Behauptung* (1). Wir zeigen hier sogleich etwas mehr. — In den Punkten von $\mathfrak{C}$ ist $F_x(t, x)$, stetig, und es existiert auf $\mathfrak{C}$ $F_x(t, x)^{-1}$. Nach Hilfssatz 1 ist dann auch $F_x(t, x)^{-1}$ auf $\mathfrak{C}$ stetig, ebenso damit $\|F_x(t, x)^{-1}\|$. Wegen der Kompaktheit von $\mathfrak{C}$ existiert also ein $m > 0$ mit

$$\|F_x(t, x)^{-1}\| \leq \frac{1}{m} \qquad ((t, x) \in \mathfrak{C}).$$

Aus der Stetigkeit von $F_x(t, x)$ in den Punkten von $\mathfrak{C}$ und der Kompaktheit von $\mathfrak{C}$ ergibt sich weiter die Existenz eines $r_1 > 0$, derart daß für jedes Paar $(t, x) \in \mathfrak{M}$, $(t_0, x_0) \in \mathfrak{C}$ mit $\delta((t, x), (t_0, x_0)) < r_1$ gilt

$$\|F_x(t, x) - F_x(t_0, x_0)\| \leq \frac{m}{2}.$$

Denn sonst hätte man Folgen $(t_n, x_n) \in \mathfrak{M}$, $(t_{0n}, x_{0n}) \in \mathfrak{C}$ mit

$$\delta((t_n, x_n), (t_{0n}, x_{0n})) \to 0, \qquad \|F_x(t_n, x_n) - F_x(t_{0n}, x_{0n})\| > \frac{m}{2}.$$

Man wählt nun eine Teilfolge der (t_{0n}, x_{0n}) aus, die gegen ein $(t_{00}, x_{00}) \in \mathfrak{C}$ konvergiert[13]). Dann folgt ein Widerspruch gegen die Stetigkeit von $F_x(t, x)$ an dieser Stelle.

Aus diesen beiden Tatsachen folgt nun mit Hilfssatz 1:

In den Punkten $(t, x) \in \mathfrak{M}$ mit $\delta((t, x), \mathfrak{C}) < r_1$ existiert $F_x(t, x)^{-1}$. Man hat $\|F_x(t, x)^{-1}\| \leq \frac{2}{m}$. $F_x(t, x)^{-1}$ ist in den Punkten von $\mathfrak{C}$ stetig.

Nun zur eigentlichen Behauptung (1). Wäre sie unrichtig, so hätte man Folgen $(t_n, x_{1n}) \in \mathfrak{M}$, $(t_n, x_{2n}) \in \mathfrak{M}$ mit $x_{1n} \neq x_{2n}$, $\delta((t_n, x_{in}), \mathfrak{C}) \to 0$, $F(t_n, x_{1n}) = F(t_n, x_{2n})$. Wegen der Kompaktheit von $\mathfrak{C}$ kann man annehmen

$$(t_n, x_{1n}) \to (t_0, x_{10}) \in \mathfrak{C},$$

$$(t_n, x_{2n}) \to (t_0, x_{20}) \in \mathfrak{C}.$$

[13]) Entsprechendes können wir als Definition der Kompaktheit wählen.

Dann folgt mit der Stetigkeit von F in den Punkten von $\mathfrak{C}$

$$F(t_0, x_{10}) = F(t_0, x_{20}),$$

also nach Voraussetzung $x_{10} = x_{20} = x_0$. Ist nun aber

$$\delta\big((t_n, x_{1n}), (t_0, x_0)\big) < r_1,$$
$$\delta\big((t_n, x_{2n}), (t_0, x_0)\big) < r_1,$$

so kann man $F(x) = F(t_n, x)$, $A = F_x(t_0, x_0)$ wählen und erhält mit der ersten Behauptung von Hilfssatz 2 einen Widerspruch. — Wir nehmen im folgenden $r \leq r_1$.

Behauptung (2). Die Eineindeutigkeit ist nach Voraussetzung gegeben. $\tilde{\mathfrak{C}}$ ist ferner kompakt wegen der Stetigkeit von F auf $\mathfrak{C}$. Denn: seien $(t_n, y_n) \in \tilde{\mathfrak{C}}$, $y_n = F(t_n, x_n)$, $(t_n, x_n) \in \mathfrak{C}$. Dann gibt es eine Teilfolge (t_{n_ν}, x_{n_ν}), die gegen ein $(t_0, x_0) \in \mathfrak{C}$ konvergiert. Wegen der Stetigkeit von F auf $\mathfrak{C}$ folgt dann $y_{n_\nu} \to y_0 = F(t_0, x_0)$, $(t_{n_\nu}, y_{n_\nu}) \to (t_0, y_0) \in \tilde{\mathfrak{C}}$.

Behauptung (3). Wäre sie nicht erfüllt, so gäbe es eine Folge (t_n, y_n) mit $\delta\big((t_n, y_n), \tilde{\mathfrak{C}}\big) \to 0$, derart daß kein x_n mit $y_n = F(t_n, x_n)$ und $\delta\big((t_n, x_n), \mathfrak{C}\big) < r$ existiert. Wir können nun wegen der Kompaktheit von $\tilde{\mathfrak{C}}$ annehmen, daß $(t_n, y_n) \to (t_0, y_0) \in \tilde{\mathfrak{C}}$ mit $y_0 = F(t_0, x_0)$, $(t_0, x_0) \in \mathfrak{C}$. Das führt zu einem Widerspruch.

Sei für $\|t_n - t_0\| < r$ gesetzt

$$F(t_n, x_0) = \tilde{y}_n.$$

Dann überdeckt nach Hilfssatz 2 die Menge der $y = F(t_n, x)$ mit $\|x - x_0\| < r$ die Kugel $\|y - \tilde{y}_n\| < \frac{m}{2} r$. [Man wählt dabei in Hilfssatz 2 $a = x_0$, $F(x) = F(t_n, x)$, $A = F_x(t_0, x_0)$, beachtet die hier besonders bequeme Wahl der Norm in $\mathfrak{B}_0 \times \mathfrak{B}_1$ [4]) und die Konstruktion von $r \leq r_1$.] Nun gilt nach Voraussetzung der Stetigkeit von F in den Punkten von $\mathfrak{C}$

$$\tilde{y}_n \to y_0.$$

Dann wird aber für $\|\tilde{y}_n - y_0\| < \delta$ von den

$$y = F(t_n, x), \quad \|x - x_0\| < r$$

die Kugel

$$\|y - y_0\| < \frac{m}{2} r - \delta$$

überdeckt, was den Widerspruch zur Konstruktion der (t_n, y_n) liefert.

Behauptung (4). Sei (t, y_1) im Definitionsbereich (*), $x_1 = \Phi(t, y_1)$. Dann gibt es nach Konstruktion von ϱ, $r \leq r_1$ ein $(t_0, x_0) \in \mathfrak{C}$ mit $\delta\big((t, x_1), (t_0, x_0)\big) < r$. Es gibt weiter ein $r_2 > 0$, derart daß für x mit $\|x - x_1\| < r_2$ noch immer $\delta\big((t, x), (t_0, x_0)\big) < r$, also $\|F_x(t, x) - F_x(t_0, x_0)\| \leq \frac{m}{2}$ gilt. Nach Hilfssatz 2 folgt dann, daß die $y = F(t, x)$ mit $\|x - x_1\| < r_2$ die Kugel $\|y - y_1\| < \frac{m}{2} r_2$

überdecken. Für y in dieser Kugel und das zugehörige $x = \Phi(t, y)$ hat man ferner

$$\frac{m}{2}\|x - x_1\| \leq \|y - y_1\|.$$

Da nun $F_x(t, x_1)^{-1}$ existiert, folgt mit

$$y - y_1 = F_x(t, x_1)(x - x_1) + o(\|x - x_1\|)$$

die Behauptung:

$$x - x_1 = F_x(t, x_1)^{-1}(y - y_1) + o(\|y - y_1\|).$$

Behauptung (5). Wie der bisherige Beweis zeigt, kann man das r beliebig klein wählen und jeweils ein entsprechendes ϱ finden. Hat man daher $(t_n, y_n) \to (t_0, y_0) \in \tilde{\mathfrak{C}}$, $y_0 = F(t_0, x_0)$, $(t_0, x_0) \in \mathfrak{C}$, so muß notwendig $\delta((t_n, x_n), \mathfrak{C}) \to 0$ für die $x_n = \Phi(t_n, y_n)$ gelten. Hier kann nun [($\mathfrak{C}$) kompakt] keine Teilfolge $(t_{n_\nu}, x_{n_\nu}) \to (t_0, \hat{x}_0)$ mit $\hat{x}_0 \neq x_0$ existieren, weil das $F(t_0, \hat{x}_0) = F(t_0, x_0) = y_0$ ergäbe. Also gilt $x_n \to x_0$, was für den Nachweis der Stetigkeit von Φ in den Punkten von $\tilde{\mathfrak{C}}$ zu zeigen war. Der Rest der Behauptung folgt nun mit

$$\Phi_y(t, y) = F_x(t, \Phi(t, y))^{-1}\ ^{16}).$$

Damit ist Satz 1 bewiesen.

3. Wir führen den Beweis für Satz 3.

Sei $a = a_1 + a_2$, $a_i \in \mathfrak{B}_i$. Dann folgt mit Anwendung von Satz 1: Es gibt ein $r > 0$ und ein $\varrho > 0$ derart, daß zu jedem $x_2 \in \mathfrak{B}_2$ mit $\|x_2 - a_2\| < \varrho$ genau ein $x_1 \in \mathfrak{B}_1$ mit $\|x_1 - a_1\| < r$ und $G(x_1 + x_2) = 0$ existiert. Wird $x_1 = \varphi(x_2)$ gesetzt, so ist $\varphi(x_2)$ stetig differenzierbar und speziell

$$\varphi'(a_2) = -G'(a)^{-1} G'(a).$$

[Man betrachtet die durch $G(x_1 + x_2) = z_1$, $x_2 = z_2$ gegebene Abbildung Γ von $\mathfrak{M} \subset \mathfrak{B}_1 \times \mathfrak{B}_2$ in $\mathfrak{B}_1 \times \mathfrak{B}_2$ um $(x_1, x_2) = (a_1, a_2)$, $(z_1, z_2) = (0, a_2)$. Für sie ist $\Gamma'(x_1, x_2)(h_1, h_2) = (G'(x)(h_1 + h_2), h_2)$, also $\Gamma'(a_1, a_2)^{-1}(k_1, k_2) = (k_1 - G'(a)^{-1} G'(a) k_2, k_2)$.]

Hieraus folgt wiederum: Es gibt ein $\hat{r} > 0$, derart daß $G(x) = 0$ bei $x = x_1 + x_2$, $x_i \in \mathfrak{B}_i$, $\|x - a\| < \hat{r}$ äquivalent ist zu $x_1 = \varphi(x_2)$. Damit bedeutet nun die Voraussetzung, daß $f(x)$ in a ein relatives Minimum unter der Nebenbedingung $G(x) = 0$ hat, nichts anderes als, daß

$$f(\varphi(x_2) + x_2) \qquad (x_2 \in \mathfrak{B}_2)$$

in $x = a_2$ ein relatives Minimum hat. Dann muß an der Stelle a_2 die Ableitung verschwinden: Es muß für alle $h_2 \in \mathfrak{B}_2$

$$f'(a)(\varphi'(a_2) h_2 + h_2) = 0$$

---

16) Damit erhält man in gewohnter Weise auch Aussagen über höhere Ableitungen von Φ, falls solche für F gegeben sind.

gelten. Das kann nun nach der obigen Formel für $\varphi'(a_2)$ auch

$$f'(a)\,h_2 = \left(f'(a)\,G'(a)^{-1}\right) G'(a)\,h_2$$

geschrieben werden. Für $h_1 \in \mathfrak{B}_1$ ist aber die Relation

$$f'(a)\,h_1 = \left(f'(a)\,G'(a)^{-1}\right) G'(a)\,h_1$$

nach der Definition von $G'(a)^{-1}$ trivial. So hat man nunmehr, wie behauptet, diese Relation für alle $h \in \mathfrak{B}$, also

$$f'(a) = \lambda\, G'(a)$$

mit

$$\lambda = f'(a)\,G'(a)^{-1}.$$

Nun zur Bedingung zweiter Ordnung. Wir beweisen sie indirekt.

Nehmen wir also ein $\tilde{h} \in \mathfrak{B}$ mit

$$G'(a)\,\tilde{h} = 0, \qquad \left(\left(f''(a) - \lambda\, G''(a)\right)\tilde{h}\right)\tilde{h} < 0$$

an. Wir zeigen zunächst die Zwischenbehauptung: Es gibt ein $\varepsilon > 0$ derart, daß zu jedem τ mit $|\tau| < \varepsilon$ ein

$$x = a + \tau\tilde{h} + o(\tau) \qquad (\tau \to 0)$$

aus $\mathfrak{M}$ mit $G(x) = 0$ existiert.

Denn sei $\tilde{h} = \tilde{h}_1 + \tilde{h}_2$, $\tilde{h}_i \in \mathfrak{B}_i$. Wir wählen dann $\varepsilon \|\tilde{h}_2\| \leq \varrho$ (vgl. den Beginn des Beweises). Nach der obigen Überlegung gibt es für τ mit $|\tau| < \varepsilon$ genau ein $h_1(\tau) \in \mathfrak{B}_1$ mit $\|h_1(\tau)\| < r$ derart, daß $G(x) = 0$ für $x = a + h_1(\tau) + \tau\tilde{h}_2$ gilt. Man hat (Definition der Differenzierbarkeit)

$$G'(a)\,h_1(\tau) + \tau\, G'(a)\,\tilde{h}_2 + o\left(\|h_1(\tau) + \tau\tilde{h}_2\|\right) = 0$$

und nach Voraussetzung

$$G'(a)\,\tilde{h}_1 + G'(a)\,\tilde{h}_2 = 0,$$

also

$$G'(a)\left(h_1(\tau) - \tau\tilde{h}_1\right) = o\left(\|h_1(\tau) + \tau\tilde{h}_2\|\right).$$

„Multiplikation" mit $G'(a)^{-1}$ liefert nun

$$h_1(\tau) - \tau\,h_1 = o\left(\|\tilde{h}_1(\tau) + \tau\tilde{h}_2\|\right).$$

Das heißt aber

$$\|h_1(\tau) - \tau\tilde{h}_1\| \leq o(1)\left\{\|h_1(\tau) - \tau\tilde{h}_1\| + |\tau|\,\|\tilde{h}\|\right\}.$$

Für $o(1) < 1$ folgt

$$\|h_1(\tau) - \tau\tilde{h}_1\| \leq \frac{o(1)}{1 - o(1)}\,|\tau|\,\|\tilde{h}\|,$$

also, wie für die Zwischenbehauptung zu zeigen war,

$$h_1(\tau) - \tau\tilde{h}_1 = o(\tau).$$

Für die konstruierten

$$x = a + \tau \tilde{h} + o(\tau)$$

wenden wir nun den Taylorschen Satz an[15]) und beachten $f'(a) - \lambda G'(a) = 0$:

$$\begin{aligned} 0 \leq f(x) - f(a) &= (f(x) - \lambda G(x)) - (f(a) - \lambda G(a)) \\ &= \tfrac{1}{2}\left[(f''(a) - \lambda G''(a))(x-a)\right](x-a) + o(\|x-a\|^2) \\ &= \tfrac{1}{2}\tau^2\left\{\left((f''(a) - \lambda G''(a))\tilde{h}\right)\tilde{h}\right\} + o(\tau^2). \end{aligned}$$

Das liefert aber für kleine τ einen Widerspruch.

Herrn ERNST HÖLDER verdanke ich wertvolle Bemerkungen, insbesondere die Literaturhinweise auf J. HADAMARD und H. HAHN in den Anmerkungen 6 und 9.

[15]) Dabei wird die Stetigkeit der zweiten Ableitungen benutzt.

Mainz, Mathematisches Institut der Universität

(Eingegangen am 17. November 1956)

EICHLER, M.
Math. Zeitschr. Bd. 67, S. 267—298 (1957)

Eine Verallgemeinerung der Abelschen Integrale

Herrn C. L. SIEGEL zum 60. Geburtstage gewidmet

Von

M. EICHLER

Einleitung

G. BOL [1] entdeckte, daß die $(n-1)$-fache Ableitung einer Modulform des (nicht negativen) Grades $n-2$ eine Modulform $(-n)$-ten Grades ist. H. PETERSSON griff diesen Gedanken in einer großen Abhandlung [5] auf, allerdings gewissermaßen nur am Rande. Angeregt durch Herrn PETERSSON, habe ich die umgekehrten Bildungen, nämlich die $(n-1)$-fachen unbestimmten Integrale von Modulformen des Grades $-n$ studiert, vorwiegend unter der Voraussetzung $n \equiv 0 \bmod 2$, und lege die Ergebnisse im folgenden vor.

Man braucht sich dabei keineswegs auf Modulfunktionen zu beschränken. Vielmehr geht man von einer Riemannschen Fläche $\mathfrak{R}$ und einer linearen Differentialgleichung 2. Ordnung auf $\mathfrak{R}$ aus. Der Quotient τ zweier Lösungen von dieser ist eine auf $\mathfrak{R}$ linear polymorphe Funktion, für welche Singularitäten bekannter Art zugelassen werden. Im einzelnen siehe zu den gemachten Voraussetzungen den § 1. Die Funktionen x und die Differentiale $x\,dy^l$ beliebigen Grades werden als automorphe Funktionen $x(\tau)$ und automorphe Formen $x\,dy^l = \varphi(\tau)\,d\tau^l$ betrachtet. Es wird nicht vorausgesetzt, daß τ eine Grenzkreisuniformisierende ist. Nunmehr lassen sich die bekannten Sätze über Abelsche Integrale, ihre Perioden und die Riemannschen Periodenrelationen übertragen (§§ 2—4). Die Analoga der Abelschen Integrale haben Perioden, welche Polynome in τ sind.

Hierauf werden die Ergebnisse auf gewisse Korrespondenzen von $\mathfrak{R}$ zu sich angewandt, die in § 5 von unserem Standpunkt aus überhaupt erst erklärt werden müssen. Die Korrespondenzen erfahren in den Moduln der Differentiale 1. Gattung Darstellungen. Die Spur für eine solche Darstellung wird in § 6 bestimmt, und dieses scheint mir das Kernstück der Arbeit zu sein. Die zum Verständnis des Ergebnisses (58), (59) oder (60) notwendigen Erläuterungen können hier leider nicht vorweggenommen werden. Im Falle der Modulfunktionen und Modularkorrespondenzen nimmt die Spurformel die Gestalt (62) an. Diese Formel wurde erstmalig von A. SELBERG auf dem International Colloquium on Zeta Functions in Bombay im Februar 1956 ohne Beweis mitgeteilt; vgl. [8]. Andere Spezialfälle von (60), für Modulformen (-4)-ten Grades bezüglich einer Kongruenzuntergruppe der Modulgruppe, waren mir seit dem Jahre 1954 bekannt; sie lassen sich alle aus dem Spezialfall (62) für $n=4$ herleiten, für den ich kürzlich einen (überaus komplizierten) Beweis veröffentlichte [3].

Man wird bei aufmerksamem Lesen bemerken, daß die wesentlichsten Schlüsse in dieser Arbeit von algebraischer Natur sind. Der hier immer wieder benutze Cauchysche Integralsatz ist in diesem Zusammenhang ja nichts anderes als der rein algebraische Satz, daß die Summe der Residuen eines Differentials 0 ist. Jedoch, wenn auch die Eliminierung wesentlich analytischer Hilfsmittel den Erfolg hatte, die Spurformel unter allgemeineren Voraussetzungen zu beweisen, so bleibt doch die reizvolle Aufgabe, den Beweis noch weiter mit algebraischen Methoden zu durchdringen. Hierbei wird man sich allerdings auf einen weiteren Standpunkt stellen müssen, den A. WEIL im Jahre 1938 aufgezeigt hat [6], [7]. Der Zusammenhang dieser Arbeit mit denen von WEIL wird ganz am Schluß auseinandergesetzt werden.

§ 1. Voraussetzungen und Bezeichnungen

1. Eine endliche Riemannsche Fläche $\mathfrak{R}$ sei vorgelegt; ihr Geschlecht sei g. Auf $\mathfrak{R}$ seien h Punkte $\mathfrak{s}_\nu$ und k Punkte $\mathfrak{e}_\lambda$ ausgezeichnet; den letzteren seien natürliche Zahlen $n_\lambda > 1$ zugeordnet. Dabei wird

$$2g - 2 + h + \sum_{\lambda=1}^{k} \left(1 - \frac{1}{n_\lambda}\right) > 0 \tag{1}$$

vorausgesetzt. Die $\mathfrak{s}_\nu$ werden stets als die *Spitzen*, die $\mathfrak{e}_\lambda$ als die *Ecken* bezeichnet werden; der Grund dieser Bezeichnungsweise wird später klar werden.

Es gebe auf der so punktierten Fläche $\mathfrak{R}$ eine Funktion τ mit folgenden drei Eigenschaften: 1) Die inverse Abbildungsfunktion der τ-Ebene auf $\mathfrak{R}$ ist überall, auch in den Ecken und Spitzen, holomorph. 2) Bei analytischer Fortsetzung längs geschlossener Wege erfährt τ lineare Transformationen $\varepsilon\tau = \frac{a\tau + b}{c\tau + d}$. 3) Bei n_λ-facher Umkreisung der Ecke $\mathfrak{e}_\lambda$ geht τ in sich über. Dasselbe gilt dann natürlich auch bei analytischer Fortsetzung längs solcher Wege, die aus n_λ-fachen Umkreisungen der $\mathfrak{e}_\lambda$ kombinierbar sind. Wir setzen nun voraus, daß τ bei Fortsetzung längs anderer geschlossener Wege nicht in sich selber übergeht.

Die Funktion τ läßt sich in bekannter Weise als Quotient zweier linear unabhängiger Lösungen einer linearen Differentialgleichung auf $\mathfrak{R}$ darstellen, deren höchster Koeffizient 1 ist, und deren übrige Koeffizienten in den Ecken und Spitzen in gewisser Weise singulär sind, aber auch nur dort.

Die Funktion τ bildet nun die Riemannsche Fläche $\mathfrak{R}$, von der man nur die Spitzen entfernt hat, auf eine einfach zusammenhängende Riemannsche Fläche $\mathfrak{T}$ ab. $\mathfrak{T}$ ist eine Überlagerung von $\mathfrak{R}$, dabei sind sämtliche über einer Ecke $\mathfrak{e}_\lambda$ liegenden Punkte von $\mathfrak{T}$ n_λ-fache Verzweigungen. In welcher Weise $\mathfrak{T}$ über der komplexen τ-Ebene liegt, interessiert hier nicht; insbesondere wird nicht vorausgesetzt, daß $\mathfrak{T}$ das Innere eines Halbkreises ist, wie bei der Grenzkreisuniformisierung. Aus der Voraussetzung (1) läßt sich allerdings folgern, daß $\mathfrak{T}$ weder die volle τ-Ebene noch die um den unendlich fernen Punkt ergänzte Vollebene ist.

Die Gruppe der geschlossenen Wege in der punktierten Fläche $\mathfrak{R}$ (mit der zusätzlichen Verabredung, daß eine n_λ-fache Umkreisung von $\mathfrak{e}_\lambda$ als ein geschlossener Weg anzusehen ist), stellt sich nun einerseits durch lineare Substitutionen dar, andererseits als eine eigentlich diskontinuierliche Gruppe konformer Abbildungen von $\mathfrak{T}$ auf sich. Ein Fundamentalbereich $\mathfrak{F}$ dieser Gruppe ist dabei ein Abbild der geeignet aufgeschnittenen Fläche $\mathfrak{R}$, welches überall konform ist, abgesehen von den Ecken $\mathfrak{e}_\lambda$, wo eine Vollumgebung in $\mathfrak{R}$ auf einen Sektor vom Öffnungswinkel $\frac{2\pi}{n_\lambda}$ in $\mathfrak{F}$ abgebildet wird, und ferner abgesehen von den Spitzen, wo eine Vollumgebung in $\mathfrak{R}$ auf einen von zwei einander berührenden Kreisen gebildeten krummlinigen Sektor in $\mathfrak{F}$ abgebildet wird. Das heißt, $\mathfrak{F}$ hat in den zu den $\mathfrak{e}_\lambda$ gehörigen Punkten Ecken vom Winkel $\frac{2\pi}{n_\lambda}$ und in den zu den $\mathfrak{s}_\nu$ gehörigen Punkten Spitzen vom Winkel 0.

Die Gruppe der Abbildungen von $\mathfrak{T}$ auf sich soll mit Γ bezeichnet werden; sie besteht, wie wir sahen, aus linearen Substitutionen. Durch Vorgabe eines Fundamentalbereichs in der bekannten kanonischen Form wird gleichzeitig ein System von Erzeugenden $\alpha_\mu, \beta_\mu, \gamma_\nu, \delta_\lambda$ ($\mu = 1, 2, \ldots, g$; $\nu = 1, 2, \ldots, h$; $\lambda = 1, 2, \ldots, k$) mit den definierenden Relationen

$$\alpha_1\beta_1\alpha_1^{-1}\beta_1^{-1}\ldots\alpha_g\beta_g\alpha_g^{-1}\beta_g^{-1}\gamma_1\ldots\gamma_h\delta_1\ldots\delta_k = \delta_\lambda^{n_\lambda} = 1 \tag{2}$$

geliefert; wir haben hier die Gruppeneins mit dem Symbol 1 bezeichnet. Die linearen Substitutionen werden stets so geschrieben werden:

$$\varepsilon\tau = \frac{a_\varepsilon\tau + b_\varepsilon}{c_\varepsilon\tau + d_\varepsilon}, \qquad \varepsilon = \begin{pmatrix} a_\varepsilon & b_\varepsilon \\ c_\varepsilon & d_\varepsilon \end{pmatrix}.$$

Über die Elemente von Γ ist noch das Folgende bekannt: 1) Die γ_ν lassen sich durch geeignete, nicht notwendig zu Γ gehörige lineare Substitutionen ϱ in die Normalform

$$\varrho^{-1}\gamma_\nu\varrho = a\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, \qquad a \neq 0 \tag{3}$$

transformieren. Sie besitzen in dieser Form den Fixpunkt $\tau = i\infty$, und dieser ist ein Randpunkt der Fläche $\mathfrak{T}$. Eine lokale Uniformisierende in der Umgebung von $\mathfrak{s}_\nu$ ist

$$z = e^{2\pi i\tau}. \tag{4}$$

2) Die δ_λ sind durch geeignete, nicht notwendig zu Γ gehörige lineare Substitutionen in die Normalform

$$\varrho^{-1}\delta_\lambda\varrho = a\begin{pmatrix} \zeta & 0 \\ 0 & 1 \end{pmatrix}, \qquad a \neq 0 \tag{5}$$

transformierbar, wo ζ eine primitive n_λ-te Einheitswurzel ist. Sie besitzen in dieser Form die Fixpunkte $\tau = 0, \infty$ und von diesen liegt einer im Inneren von $\mathfrak{T}$. Eine lokale Uniformisierende in der Umgebung von $\mathfrak{e}_\lambda$ ist

$$z = \tau^{n_\lambda}. \tag{6}$$

3) Alle Elemente ε, welche nicht mit einem der γ_ν oder δ_λ konjugiert sind, lassen sich durch geeignete, nicht notwendig zu Γ gehörige lineare Substitutionen in die Normalform

$$\varrho^{-1}\varepsilon\varrho = a\begin{pmatrix}\vartheta & 0\\ 0 & \vartheta^{-1}\end{pmatrix}, \qquad a \neq 0 \tag{7}$$

mit reellem $\vartheta \neq 1$ transformieren. Die Fixpunkte eines solchen ε gehören nicht zu $\mathfrak{T}$.

2. Es sei $n \geq 1$ eine bis zum Schluß festgehaltene natürliche Zahl. Wir betrachten homogene Funktionen $\varphi(\tau_1, \tau_2)$ in zwei Variablen vom Grade $-n$ und homogene Funktionen $\Phi(\tau_1, \tau_2)$ vom Grade $(n-2)$. Durch die Festsetzungen

$$\varphi(\tau_1, \tau_2)\,[\varepsilon]^{-n} = \varphi(a_\varepsilon\tau_1 + b_\varepsilon\tau_2, c_\varepsilon\tau_1 + d_\varepsilon\tau_2)\cdot(a_\varepsilon d_\varepsilon - b_\varepsilon c_\varepsilon)^{\frac{n}{2}},$$
$$\Phi(\tau_1, \tau_2)\,[\varepsilon]^{n-2} = \Phi(a_\varepsilon\tau_1 + b_\varepsilon\tau_2, c_\varepsilon\tau_1 + d_\varepsilon\tau_2)\cdot(a_\varepsilon d_\varepsilon - b_\varepsilon c_\varepsilon)^{\frac{2-n}{2}}$$

werden jeder linearen Substitution ε je ein Operator $[\varepsilon]^{-n}$ und $[\varepsilon]^{n-2}$ im Raume der homogenen Funktionen der Grade $-n$ und $n-2$ zugeordnet. Speziell ist $[1]^{-n}$ und $[1]^{n-2}$ jeweils der Einheitsoperator. Wir werden allerdings die zugehörigen inhomogenen Formen

$$\varphi(\tau) = \varphi\left(\frac{\tau_1}{\tau_2}\right) = \varphi(\tau_1, \tau_2)\,\tau_2^n, \qquad \Phi(\tau) = \Phi\left(\frac{\tau_1}{\tau_2}\right) = \Phi(\tau_1, \tau_2)\,\tau_2^{2-n}$$

benutzen. Dann schreiben sich obige Operatoren so:

$$\varphi(\tau)\,[\varepsilon]^{-n} = \varphi(\varepsilon\tau)\,\frac{(a_\varepsilon d_\varepsilon - b_\varepsilon c_\varepsilon)^{\frac{n}{2}}}{(c_\varepsilon\tau + d_\varepsilon)^n}, \tag{8}$$

$$\Phi(\tau)\,[\varepsilon]^{n-2} = \Phi(\varepsilon\tau)\,\frac{(c_\varepsilon\tau + d_\varepsilon)^{n-2}}{(a_\varepsilon d_\varepsilon - b_\varepsilon c_\varepsilon)^{\frac{n-2}{2}}}. \tag{9}$$

Bei ungeradem n ist das Vorzeichen in (8) und (9) zweideutig. Wir werden diese Operatoren aber nur dann bei ungeradem n benutzen, wenn $a_\varepsilon d_\varepsilon - b_\varepsilon c_\varepsilon > 0$ ist und verabreden, daß auch die Quadratwurzel positiv genommen werden soll. Jetzt sind (8) und (9) eindeutig definiert.

Die Bezeichnung von Formen vom Grade $-n$ mit kleinen griechischen Buchstaben und von Formen vom Grade $n-2$ mit großen griechischen Buchstaben wird durchweg festgehalten.

Spezielle Formen vom Grade $n-2$ sind die Polynome dieses Grades. Für sie wird durch (9) eine Darstellung der projektiven linearen Gruppe in zwei Variablen geliefert.

3. Unter einer *automorphen Form* $\varphi(\tau)$ *vom Grade* $-n$ versteht man eine Funktion, die den Funktionalgleichungen $\varphi(\tau)\,[\varepsilon]^{-n} = \varphi(\tau)$ für alle $\varepsilon \in \Gamma$ genügt, und die überall auf $\mathfrak{R}$, einschließlich der Ecken und Spitzen meromorph ist. Entsprechendes verlangt man von einer automorphen Form $\Phi(\tau)$ vom Grade $n-2$.

Wegen

$$d\varepsilon\tau = \frac{a_\varepsilon d_\varepsilon - b_\varepsilon c_\varepsilon}{(c_\varepsilon \tau + d_\varepsilon)^2}\, d\tau$$

ist für eine automorphe Form $\varphi(\tau)$ von geradem Grade $-n = -2f$: $\varphi(\tau)\, d\tau^f$ ein invariantes Differential f-ten Grades, d.h. es gilt

$$\varphi(\varepsilon\tau)\, d\varepsilon\tau^f = \varphi(\tau)\, d\tau^f.$$

Einem solchen Differential ist bekanntlich durch die folgende Vorschrift ein Divisor $\mathfrak{f}$ zugeordnet: $\mathfrak{f}$ ist durch einen Primdivisor $\mathfrak{p}$ ebensooft teilbar wie die Funktion $\varphi(\tau)\left(\frac{d\tau}{dz}\right)^f$, wo z eine lokale Uniformisierende an der Stelle $\mathfrak{p}$ (die mit dem gleichen Symbol wie der zugehörige Primdivisor bezeichnet werden darf) ist. Dieser Divisor gehört der f-ten Potenz der Differentialklasse W an. Umgekehrt läßt sich auch zu jedem Divisor $\mathfrak{f}$ aus dieser Klasse eine automorphe Form des Grades $-n$ angeben.

Wir wollen nun die Dimension der Schar der *ganzen automorphen Spitzenformen* des Grades $-n = -2f$ bestimmen. Darunter versteht man die überall auf $\mathfrak{R}$ holomorphen und in den Spitzen verschwindenden automorphen Formen. Eine in einem gewöhnlichen Punkt $\mathfrak{p}$ holomorphe Form ist dadurch gekennzeichnet, daß der zugehörige Divisor $\mathfrak{f}$ für $\mathfrak{p}$ ganz ist, denn man kann als lokale Uniformisierende $z = \tau$ nehmen. In einer Ecke $\mathfrak{e}_\lambda$ nehmen wir die Uniformisierende (6). Es ist dann

$$\frac{d\tau}{dz} = \frac{1}{n_\lambda} z^{\frac{1}{n_\lambda} - 1}.$$

Demnach ist die in $\mathfrak{f}$ enthaltene Potenz von $\mathfrak{e}_\lambda$ unter der Voraussetzung, daß $\varphi(\tau)$ hier holomorph ist, mindestens $\mathfrak{e}_\lambda^{-\left[\left(1-\frac{1}{n_\lambda}\right)f\right]}$. In einer Spitze $\mathfrak{s}_\nu$ nehmen wir die Uniformisierende (4). Es ist dann

$$\frac{d\tau}{dz} = \frac{1}{2\pi i} z^{-1}.$$

Demnach ist die in $\mathfrak{f}$ enthaltene Potenz von $\mathfrak{s}_\nu$ unter der Voraussetzung, daß $\varphi(\tau)$ hier verschwindet, mindestens $\mathfrak{s}_\nu^{1-f}$.

Der Divisor $\mathfrak{f}$ hat hiernach die Form

$$\mathfrak{f} = \mathfrak{g} \cdot \prod_{\nu=1}^{h} \mathfrak{s}_\nu^{1-f} \cdot \prod_{\lambda=1}^{k} \mathfrak{e}_\lambda^{-\left[\left(1-\frac{1}{n_\lambda}\right)f\right]} = \mathfrak{g} \cdot \mathfrak{P}^{-1}$$

mit einem ganzen Divisor $\mathfrak{g}$. Dieser gehört der Klasse $W^f\mathfrak{P}$ an, deren Grad $2f(g-1) + h(f-1) + \sum\left[\left(1-\frac{1}{n_\lambda}\right)f\right]$ ist. Die Anzahl der linear unabhängigen ganzen Divisoren $\mathfrak{g}$ dieser Klasse ist nach dem Riemann-Rochschen Satze

$$G_n = (n-1)(g-1) + h\left(\frac{n}{2} - 1\right) + \sum_{\lambda=1}^{k}\left[\left(1-\frac{1}{n_\lambda}\right)f\right] + \dim\left(W^{1-f}\mathfrak{P}^{-1}\right)$$

Die Klasse $W^{1-f}\mathfrak{P}^{-1}$ hat den Grad

$$\operatorname{Grad}(W^{1-f}\mathfrak{P}^{-1}) = -(f-1)(2g-2+h) - \sum_{\lambda=1}^{h}\left[\left(1-\frac{1}{n_\lambda}\right)f\right].$$

Dieser ist 0 für $f=1$; dann ist $W^{1-f}\mathfrak{P}^{-1}$ die Einheitsklasse, und ihre Dimension gleich 1. Wir behaupten, daß $\operatorname{Grad}(W^{1-f}\mathfrak{P}^{-1})<0$ für $f>1$ ist. Zum Beweise verifiziert man zunächst

$$\left[\left(1-\frac{1}{n_\lambda}\right)f\right] = f-1-\left[\frac{f-1}{n_\lambda}\right]. \tag{10}$$

Jetzt ist

$$\frac{1}{f-1}\operatorname{Grad}(W^{1-f}\mathfrak{P}^{-1})$$
$$= -\left(2g-2+h+\sum_{\lambda=1}^{h}\left(1-\frac{1}{f-1}\left[\frac{f-1}{n_\lambda}\right]\right)\right) \leq -\left(2g-2+h+\sum_{\lambda=1}^{h}\left(1-\frac{1}{n_\lambda}\right)\right),$$

und das ist <0 wegen der Voraussetzung (1).

Das Endergebnis lautet: *Der Rang der Schar der ganzen automorphen Spitzenformen von geradem Grade $-n$ ist unter der Voraussetzung* (1)

$$G_n = \begin{cases} g & \text{für } n=2, \\ (n-1)(g-1)+h\dfrac{n-2}{2}+\displaystyle\sum_{\lambda=1}^{h}\left(\frac{n-2}{2}-\left[\frac{n-2}{2n_\lambda}\right]\right) & \text{für } n>2. \end{cases} \tag{11}$$

§ 2. Die Integrale und ihre Perioden

1. Eine ganze automorphe Spitzenform vom Grade $-n$ soll fortan kurz eine *automorphe Form $(-n)$-ten Grades und 1. Gattung* genannt werden. Für eine solche ist das Integral

$$\Phi(\tau) = \frac{1}{(n-2)!}\int_{\tau_0}^{\tau}(x-\sigma)^{n-2}\varphi(\sigma)\,d\sigma$$

vom Wege unabhängig. Eine nicht überall holomorphe Spitzenform (die also in allen Spitzen verschwindet), für die aber dieses Integral auch wegunabhängig ist, soll eine *automorphe Form $(-n)$-ten Grades und 2. Gattung* heißen. Ein solches $\varphi(\tau)$ hat in dem Pol $\tau=\tau_1$ den Hauptteil

$$\sum_{m\geq 0}\frac{c_m}{(\tau-\tau_1)^{n+m}}.$$

Die Integrale $\Phi(\tau)$ heißen *Integrale 1. oder 2. Gattung*. Sie können auch durch $(n-1)$-fache unbestimmte Integration der Formen $(-n)$-ten Grades gewonnen werden. Es ist daher sinngemäß, ein willkürliches Polynom $\Theta(\tau)$ vom Grade $n-2$ als „Integrationskonstante" auf der rechten Seite hinzuzufügen:

$$\Phi(\tau) = \frac{1}{(n-2)!}\int_{\tau_0}^{\tau}(\tau-\sigma)^{n-2}\varphi(\sigma)\,d\sigma + \Theta(\tau). \tag{12}$$

Durch die Variablensubstitution $\sigma = \varepsilon(\zeta)$ auf der rechten Seite von (12) zeigt man

$$\Phi(\tau)\,[\varepsilon]^{n-2} = \Phi(\tau) + \Omega_\varepsilon(\tau) \tag{13}$$

mit

$$\begin{cases} \Omega_\varepsilon(\tau) = \Theta(\tau)\left([\varepsilon]^{n-2} - [1]^{n-2}\right) + \\ \qquad + \frac{1}{(n-2)!}\int\limits_{\tau_\varepsilon}^{\varepsilon\tau_0} (a_\varepsilon\tau + b_\varepsilon - (c_\varepsilon\tau + d_\varepsilon)\,\sigma)^{n-2}\,\varphi(\sigma)\,d\sigma. \end{cases} \tag{14}$$

Bei Anwendung des Operators $[\varepsilon]^{n-2}$ *addiert sich also zu* $\Phi(\tau)$ *die „Periode"* $\Omega_\varepsilon(\tau)$, *ein Polynom vom Grade* $n-2$.

2. Sei umgekehrt $\Phi(\tau)$ irgendeine meromorphe Funktion, welche für alle $\varepsilon \in \Gamma$ der Funktionalgleichung (13) mit irgendwelchen Polynomen $\Omega_\varepsilon(\tau)$ vom Grade $n-2$ genügt. Dann ist

$$\varphi(\tau) = \left(\frac{d}{d\tau}\right)^{n-1}\Phi(\tau) = \frac{(n-1)!}{2\pi i}\int \frac{\Phi(\sigma)\,d\sigma}{(\sigma-\tau)^n}, \tag{15}$$

das Integral erstreckt über einen kleinen Kreis um $\sigma = \tau$ [unter der Voraussetzung, daß $\Phi(\sigma)$ hier holomorph ist], eine automorphe Form vom Grade $-n$. Substituiert man nämlich in dem Integral $\sigma = \varepsilon(\zeta)$, so resultiert wegen (13)

$$\varphi(\tau)\,[\varepsilon]^{-n} = \frac{(n-1)!}{2\pi i}\int \frac{\Phi(\zeta)\,[\varepsilon]^{n-2}\,d\zeta}{(\zeta-\tau)^n} = \varphi(\tau).$$

Insbesondere ist die $(n-1)$-te Ableitung einer automorphen Form $(n-2)$-ten Grades eine solche vom Grade $-n$.

3. Man schreibe (13) in der Form

$$\Omega_\varepsilon = \Phi\left([\varepsilon]^{n-2} - [1]^{n-2}\right);$$

die Variable τ braucht in diesem Zusammenhang nicht erwähnt zu werden. Dann ergibt sich sofort für zwei Elemente ε, η aus Γ:

$$\Omega_{\varepsilon\eta} - \Omega_\varepsilon[\eta]^{n-2} - \Omega_\eta = 0. \tag{16}$$

Diese Gleichung läßt die folgende Deutung zu:

Die Polynome $(n-2)$-ten Grades bilden nach § 1, Nr. 2 einen Darstellungsmodul $\mathfrak{M}^{n-2}$ für die Gruppe Γ. Ω_ε ist demnach eine Abbildung von Γ in $\mathfrak{M}^{n-2}$. Man nennt allgemein Abbildungen dieser Art *Coketten*, Abbildungen der Eigenschaft (16) *geschlossene Coketten* oder *Cocyclen*. Spezielle Cocyclen sind die *Coränder* $\Theta\left([\varepsilon]^{n-2} - [1]^{n-2}\right)$. Cocyclen und Coränder bilden je eine additive Gruppe. Die Klassen der Cocyclen modulo der Untergruppe der Coränder sind die *Cohomologieklassen*.

Nach Nr. 1 ist also jeder automorphen Form 1. oder 2. Gattung eine Cohomologieklasse zugeordnet. Erst später wird sich zeigen, daß diese Zuordnung umkehrbar ist.

Wir notieren zum Schluß dieses Abschnittes zwei wichtige Eigenschaften der Perioden: *Für die speziellen Erzeugenden γ_ν, δ_λ von Γ (vgl. § 1, Nr. 1) gilt*

$$\Omega_{\gamma_\nu} = \Theta_\nu\left([\gamma_\nu]^{n-2} - [1]^{n-2}\right), \tag{17}$$

$$\Omega_{\delta_\lambda} = \Theta'_\lambda\left([\delta_\lambda]^{n-2} - [1]^{n-2}\right) \tag{18}$$

mit gewissen Polynomen $\Theta_\nu, \Theta'_\lambda$ $(n-2)$-ten Grades. Die Gl. (17) erhält man unmittelbar aus (14), wenn man als untere Grenze τ_0 des Integrals (12) die Spitze $\mathfrak{s}_\nu$ nimmt. (12) wird dann ein uneigentliches Integral, welches dann und nur dann existiert, wenn $\varphi(\tau)$ eine Spitzenform ist. Zum Beweise von (18) geht man von $\delta_\lambda^{n_\lambda} = 1$ aus. Aus (16) folgt dann

$$0 = \Omega_1 = \Omega_{\delta_\lambda^{n_\lambda}} = \Omega_{\delta_\lambda}\left([1]^{n-2} + [\delta_\lambda]^{n-2} + \cdots + [\delta_\lambda^{n_\lambda-1}]^{n-2}\right)$$

und hieraus weiter (18) mit

$$\Theta'_\lambda = \frac{1}{n_\lambda - 1}\,\Omega_{\delta_\lambda}\left((n_\lambda - 1)\,[1]^{n-2} + (n_\lambda - 2)\,[\delta_\lambda]^{n-2} + \cdots + [\delta_\lambda^{n_\lambda-1}]^{n-2}\right).$$

4. Es soll jetzt der Rang R_n der Schar der Cohomologieklassen unter der Voraussetzung (17) für alle Erzeugenden γ_ν berechnet werden. Die Zahl n soll dabei als gerade vorausgesetzt werden. Es ist

$$R_n = r_n - r_n^0, \tag{19}$$

wobei r_n, r_n^0 die Ränge der Gruppen der Cocyclen und der Coränder sind. Offenbar ist

$$r_n^0 = \begin{cases} 0 & \text{für } n = 2 \\ n-1 & \text{für } n > 2. \end{cases} \tag{20}$$

Die Berechnung von r_n verläuft folgendermaßen. Man gebe Ω_{ε_i} für alle ε_i eines Erzeugendensystems von Γ vor und kann dann Ω_ε überall in Γ unter Benutzung von (16) berechnen. Man muß bei der Vorgabe der Ω_{ε_i} allerdings darauf achten, daß für jede der definierenden Relationen $\varepsilon_1^{a_1}\varepsilon_2^{a_2}\cdots = 1$ auch

$$0 = \Omega_1 = \Omega_{\varepsilon_1^{a_1}\varepsilon_2^{a_2}\ldots} = \Omega_{\varepsilon_1}\left([1]^{n-2} + [\varepsilon_1]^{n-2} + \cdots + [\varepsilon_1^{a_1-1}]^{n-2}\right)[\varepsilon_2^{a_2}\ldots]^{n-2} + \Omega_{\varepsilon_1^{a_1}\ldots}$$

gilt. Es wird das Erzeugendensystem aus § 1, Nr. 1 genommen werden.

Zunächst sei $n = 2$. Aus (17) folgt $\Omega_{\gamma_\nu} = 0$. Aus (18) folgt $\Omega_{\delta_\lambda} = 0$. Damit sind bereits alle Relationen (2) erfüllt bis auf die erste. Für $n = 2$ schreibt sich (16) so: $\Omega_{\varepsilon\eta} = \Omega_\varepsilon + \Omega_\eta$. Man kann die $\Omega_{\alpha_\mu}, \Omega_{\beta_\mu}$ willkürlich vorgeben, die erste Relation (2) ist erfüllt:

$$\Omega_{\alpha_1\beta_1\alpha_1^{-1}\beta_1^{-1}\ldots} = \Omega_{\alpha_1} + \Omega_{\beta_1} - \Omega_{\alpha_1} - \Omega_{\beta_1} + - \cdots = 0.$$

Also ist $R_2 = g$.

Von jetzt ab sei $n > 2$. Wir unterscheiden verschiedene Fälle; zuerst werde $g > 0$ angenommen. Man gebe alle $\Omega_{\alpha_\mu}, \Omega_{\beta_\mu}$ bis auf $\Omega_{\alpha_1}, \Omega_{\beta_1}$ willkürlich vor, ebenso setze man die $\Omega_{\gamma_\nu}, \Omega_{\delta_\lambda}$ in der Form (17), (18) mit willkürlichen $\Theta_\nu, \Theta'_\lambda$ an, und suche nun $\Omega_{\alpha_1}, \Omega_{\beta_1}$ so zu bestimmen, daß

$$\Omega_{\alpha_1\beta_1\alpha_1^{-1}\beta_1^{-1}}[\alpha_2\beta_2\ldots]^{n-2} + \Omega_{\alpha_2\beta_2\ldots} = 0 \tag{21}$$

gilt. (21) entspricht der ersten Relation in (2), während die übrigen bereits durch den Ansatz (18) erfüllt werden. Unter Benutzung von (16) findet man

$$\Omega_{\alpha_1^{-1}} = -\Omega_{\alpha_1}[\alpha_1^{-1}]^{n-2}$$

und weiter

$$\Omega_{\alpha_1\beta_1\alpha_1^{-1}\beta_1^{-1}} = \left(\Omega_{\alpha_1}([\beta_1]^{n-2} - [1]^{n-2}) - \Omega_{\beta_1}([\alpha_1]^{n-2} - [1]^{n-2})\right)[\alpha_1^{-1}\beta_1^{-1}]^{n-2}.$$

Demnach läßt sich (21) so schreiben:

$$(22)\quad \Omega_{\alpha_1} N_{\beta_1} - \Omega_{\beta_1} N_{\alpha_1} = C, \quad N_{\beta_1} = [\beta_1]^{n-2} - [1]^{n-2}, \quad N_{\alpha_1} = [\alpha_1]^{n-2} - [1]^{n-2},$$

ein lineares Gleichungssystem für die unbekannten Koeffizienten der Polynome $\Omega_{\alpha_1}, \Omega_{\beta_1}$, dessen rechte Seite sich aus den $\Omega_{\alpha_2}, \Omega_{\beta_2}$ usw. berechnet.

Zur Lösung von (22) gehen wir folgendermaßen vor: Wir zeigen, daß es eine Matrix P so gibt, daß

$$(23)\qquad P N_{\beta_1} + N_{\alpha_1}$$

die Determinante $\neq 0$ hat (die Operatoren $[\beta_1]^{n-2}$ usw. im Raume der Polynome $(n-2)$-ten Grades können ersichtlich auch als $(n-1)$-reihige Matrizen aufgefaßt werden). Nun schreiben wir (22) in der Form

$$\Omega_{\alpha_1} N_{\beta_1} - \Omega_{\beta_1} N_{\alpha_1} = (\Omega_{\alpha_1} + \Omega_{\beta_1} P) N_{\beta_1} - \Omega_{\beta_1}(P N_{\beta_1} + N_{\alpha_1}) = C.$$

Hier können wir $\Omega_{\alpha_1} + \Omega_{\beta_1} P$ willkürlich vorgeben und dann Ω_{β_1} und Ω_{α_1} in eindeutiger Weise berechnen. Zugleich erkennt man, daß (22) genau $n-1$ linear unabhängige Lösungen besitzt.

Zur Untersuchung der Matrix (23) nimmt man β_1 in der Normalform (7) an. Dann ist N_{β_1} die Diagonalmatrix mit den Elementen $\vartheta^{n-2}-1, \ldots, \vartheta^2-1, 0, \vartheta^{-2}-1, \ldots, \vartheta^{2-n}-1$. Es gibt also eine nicht singuläre Matrix M, so daß

$$M N_{\beta_1} = \begin{pmatrix} 1 & \cdots & 0 & 0 \\ \vdots & & \vdots & \vdots \\ 0 & \cdots & 1 & 0 \\ 0 & \cdots & 0 & 0 \end{pmatrix}$$

ist. Man schreibe $PM^{-1} = Q = (q_{ik})$, $N_{\alpha_1} = (h_{ik})$. Dann ist

$$P N_{\beta_1} + N_{\alpha_1} = \begin{pmatrix} q_{11} + h_{11} \cdots q_{1,t-1} + h_{1,t-1} & h_{1t} \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ q_{t1} + h_{t1} \cdots q_{t,t-1} + h_{t,t-1} & h_{tt} \end{pmatrix}$$

$(t = n-1)$. Wenn nun die Determinante von (23) für alle P verschwinden würde, so wäre $h_{1t} = \cdots = h_{tt} = 0$. Dann gäbe es einen sowohl von N_{β_1} wie von N_{α_1} annullierten Vektor, nämlich $(0, \ldots, 0, 1)$. Dieses führt aber auf einen Widerspruch. Dem von N_{β_1} und N_{α_1} annullierten Vektor entspricht ein bei $[\alpha_1]^{n-2}$ und $[\beta_1]^{n-2}$ invariantes Polynom $F(\tau)$. Nimmt man β_1 wie bisher in der Normalform (7) an, so ist $F(\tau) = c\,\tau^{\frac{n-2}{2}}$. Nun hat man

$$\tau^{\frac{n-2}{2}}\,[\alpha_1]^{n-2} = \left(\frac{(a_{\alpha_1}\tau + b_{\alpha_1})(c_{\alpha_1}\tau + d_{\alpha_1})}{a_{\alpha_1} d_{\alpha_1} - b_{\alpha_1} c_{\alpha_1}}\right)^{\frac{n-2}{2}} = \tau^{\frac{n-2}{2}}.$$

Das ist aber nur möglich, wenn $b_{\alpha_1} = c_{\alpha_1} = 0$ oder $a_{\alpha_1} = d_{\alpha_1} = 0$ ist. Im ersteren Falle wäre $\alpha_1 \beta_1 \alpha_1^{-1} \beta_1^{-1} = 1$, im letzteren $\alpha_1 \beta_1 \alpha_1^{-1} \beta_1^{-1} = \beta_1^{-2}$. Es würden also in beiden Fällen Relationen bestehen, die in (2) nicht vorkommen. (Wegen (1) ist nicht gleichzeitig $g = 1$, $h = k = 0$.)

Um die Gesamtheit der linear unabhängigen Cocyclen übersehen zu können, muß man noch wissen, wieviele linear unabhängige Θ_ν, Θ'_λ durch die Ansätze (17), (18) annulliert werden. Man nimmt γ_ν in der Normalform (3) an und sieht dann unmittelbar, daß $[\gamma_\nu]^{n-2} - [1]^{n-2}$ den Rang $n-2$ hat. Von $[\gamma_\nu]^{n-2} - [1]^{n-2}$ wird also nur ein Θ_ν und dessen Vielfache annulliert. Man nimmt weiter δ_λ in der Normalform (5) an. $[\delta_\lambda]^{n-2} - [1]^{n-2}$ läßt sich dann auf die Diagonalgestalt mit den Elementen $\zeta^{\frac{n-2}{2}} - 1$, $\zeta^{\frac{n-4}{2}} - 1, \ldots, \zeta - 1, 0$, $\zeta^{-1} - 1, \ldots, \zeta^{\frac{2-n}{2}} - 1$ transformieren. Von diesen sind $1 + 2\left[\frac{n-2}{2n_\lambda}\right]$ gleich 0. Ebenso viele linear unabhängige Θ'_λ gibt es, die von $[\delta_\lambda]^{n-2} - [1]^{n-2}$ annulliert werden.

Unter der Voraussetzung $g > 0$ haben wir also das Resultat erhalten:

Der Rang der Gruppe der Cohomologieklassen, deren Elemente die Eigenschaft (17) *haben, ist für ein gerades* n*:*

$$R_n = \begin{cases} 2g & (n = 2) \\ 2(n-1)(g-1) + h(n-2) + \sum_{\lambda=1}^{k} \left(n - 2 - 2\left[\frac{n-2}{2n_\lambda}\right]\right) & (n > 2). \end{cases} \tag{24}$$

Es gibt in jeder Cohomologieklasse Cocyclen, deren Koeffizienten demselben Körper angehören wie die Koeffizienten der Darstellung $[\varepsilon]^{n-2}$ *von* Γ.

Die Voraussetzung $g > 0$ läßt sich aber noch als entbehrlich nachweisen.

5. Wir beweisen (24) noch unter der Voraussetzung $g = 0$, $h > 0$. Jetzt setzt man Ω_{γ_ν} $(\nu = 2, \ldots, h)$, Ω_{δ_λ} $(\lambda = 1, \ldots, k)$ gemäß (17), (18) an und hat Ω_{γ_1} aus

$$\Omega_{\gamma_1}[\gamma_2 \ldots]^{n-2} + \Omega_{\gamma_1 \ldots} = 0 \quad \text{oder} \quad \Omega_{\gamma_1} = -\Omega_{\gamma_1 \ldots}[(\gamma_2 \ldots)^{-1}]^{n-2}$$

zu bestimmen. Durch diese Gleichung ist Ω_{γ_1} eindeutig festgelegt. Es gibt somit

$$(h-1)(n-2) + \sum_{\lambda=1}^{k} \left(n - 2 - 2\left[\frac{n-2}{2n_\lambda}\right]\right)$$

linear unabhängige Cocyclen. Diese erfüllen für $\nu = 2, \ldots, h$ die Gl. (17), dagegen nicht alle für $\nu = 1$. Setzt man γ_1 in der Normalform (3) an, so bedeutet (17), daß Ω_{γ_1} ein Polynom vom Grade $n-3$ ist. Sind Ω_ε und Ω'_ε zwei beliebige Cocyclen, so hat stets eine nicht verschwindende Linearkombination $c\Omega_\varepsilon + c'\Omega'_\varepsilon = \Omega''_\varepsilon$ die Eigenschaft, daß Ω''_{γ_1} vom Grade $n-3$ ist. Die Anzahl der linear unabhängigen Cocyclen, die für alle ν die Gl. (17) befriedigen, ist demnach um 1 kleiner als die oben angegebene Zahl. Subtrahiert man weiter

die Anzahl $r_n^0 = n-1$ der linear unabhängigen Coränder, so erhält man wieder die Anzahl (24) linear unabhängiger Cohomologieklassen der Eigenschaft (17).

Falls $g = h = 0$ ist, wird sich (24) später (in § 4) nebenher ergeben.

§ 3. Die Periodenrelationen

1. Es seien $\varphi(\tau)$ und $\psi(\tau)$ zwei automorphe Formen 1. oder 2. Gattung und $\Phi(\tau)$ das Integral (12) der ersteren. Man denke sich ferner $\mathfrak{R}$ in ein oder mehrere einfach zusammenhängende Stücke $\mathfrak{R}_i$ zerschnitten, wobei die Ecken und Spitzen auf dem Rande liegen, dagegen keine Singularitäten von $\varphi(\tau)$ oder $\psi(\tau)$. Man bilde nun die folgenden Integrale um die Ränder der $\mathfrak{R}_i$ in positivem Sinne. Die Summe dieser Integrale wird nach Division durch $2\pi i$ mit

$$I(\varphi,\psi) = \frac{1}{2\pi i}\int_{\mathfrak{R}} \Phi(\tau)\,\psi(\tau)\,d\tau \tag{25}$$

bezeichnet. *$I(\varphi,\psi)$ ist von der willkürlichen Integrationskonstanten in $\Phi(\tau)$ und von der Art der Zerschneidung von $\mathfrak{R}$ unabhängig.*

Da $\Phi(\tau)$ das $(n-1)$-fache unbestimmte Integral von $\varphi(\tau)$ ist, kann man $I(\varphi,\psi)$ durch $(n-1)$-fache partielle Integration umformen und erhält

$$I(\varphi,\psi) = (-1)^{n-1} I(\psi,\varphi). \tag{26}$$

Bei dieser Umformung zeigt es sich insbesondere, daß die Integrationskonstante von $\Phi(\tau)$ herausfällt. Also ist (25) eine Invariante von $\varphi(\tau)$ und $\psi(\tau)$.

Dieselbe Überlegung ist für die Randintegrale um die einzelnen Stücke $\mathfrak{R}_i$ von $\mathfrak{R}$ gültig. Wenn man nun von $\mathfrak{R}_i$ ein Stück $\mathfrak{R}_{i1}$ abtrennt und an ein anderes Stück $\mathfrak{R}_k$ ansetzt, so vermindert sich das Integral über $\mathfrak{R}_i$ um das über $\mathfrak{R}_{i1}$, und das Integral über $\mathfrak{R}_k$ vermehrt sich um das über $\mathfrak{R}_{i1}$, wobei der Integrand möglicherweise auf einem nicht nullhomologen Wege fortgesetzt wurde und daher um eine Periode von $\Phi(\tau)$ mal $\psi(\tau)$ anwächst. Wie wir aber schon sahen, ändert sich dabei das Integral über $\mathfrak{R}_{i1}$ nicht. Also bleibt auch die Summe (25) aller Teilintegrale bei Umformungen der Zerschneidung ungeändert.

Die Definition (25) *versagt, wenn $\varphi(\tau)$ oder $\psi(\tau)$ Singularitäten in den Ecken oder Spitzen haben.* Eine Zerschneidung von $\mathfrak{R}$ führt nämlich immer in diese hinein; daher kann man das Schnittsystem jetzt nicht so legen, daß es Singularitäten vermeidet. *In diesem Falle spare man diese Singularitäten durch kleine Kreisbögen aus, addiere dafür aber das Residuum des Integranden (gemessen in der lokalen Uniformisierenden).* Dieses Residuum liegt trotz der Willkür bei der Wahl der Integrationskonstanten von $\Phi(\tau)$ eindeutig fest, da $\psi(\tau)$ als automorphe Form 2. Gattung keine Terme $c_\nu(z-z_0)^{-\nu}$ mit $1 \leq \nu \leq n-1$ in der Laurent-Entwicklung um einen Punkt $\tau = \tau_0$ und keinen konstanten Term in der zu einer Spitze $\mathfrak{s}_\nu$ gehörigen Fourier-Entwicklung besitzt.

2. *$I(\varphi,\psi)$ verschwindet, wenn $\varphi(\tau)$ die $(n-1)$-te Ableitung einer automorphen Spitzenform $\Phi(\tau)$ vom Grade $n-2$ ist.* Dann ist nämlich der Integrand von (25) ein Differential auf $\mathfrak{R}$, welches in den Spitzen ganz ist. $I(\varphi,\psi)$ ist die Summe aller Residuen dieses Differentials, und diese ist bekanntlich 0.

$I(\varphi, \psi)$ verschwindet, wenn $\varphi(\tau)$ und $\psi(\tau)$ beide von 1. Gattung sind. Dann ist nämlich der Integrand von (25) holomorph.

3. *$I(\varphi, \psi)$ ist eine bilineare Funktion in den Perioden der Integrale von $\varphi(\tau)$ und $\psi(\tau)$.*

Sind speziell weder Ecken noch Spitzen vorhanden, so hat diese bilineare Funktion ganze rationale Koeffizienten, bis auf einen gemeinsamen Faktor $\frac{1}{2\pi i}$.

Man gehe zum Beweise von einer Zerschneidung von $\mathfrak{R}$ zu einem einfach zusammenhängenden Flächenstück aus. Der Rand besteht dann aus einer Anzahl von Schnitten A_μ, welche durch gewisse Gruppenelemente ε_μ paarweise einander zugeordnet sind; wegen der Orientierbarkeit von $\mathfrak{R}$ werden die Integrale über A_μ und $\varepsilon_\mu A_\mu$ in entgegengesetztem Sinne durchlaufen. Man hat also

$$I(\varphi, \psi) = \frac{1}{2\pi i} \sum_\mu \left(\int_{A_\mu} \Phi(\tau)\, \psi(\tau)\, d\tau - \int_{\varepsilon_\mu A_\mu} \Phi(\tau)\, \psi(\tau)\, d\tau \right).$$

Man nehme in den Integralen über $\varepsilon_\mu A_\mu$ die Variablensubstitutionen $\tau = \varepsilon_\mu \sigma$ vor und erhält dann auf Grund von (13)

$$I(\varphi, \psi) = -\frac{1}{2\pi i} \sum_\mu \int_{A_\mu} \Omega_{\varepsilon_\mu}(\tau)\, \psi(\tau)\, d\tau,$$

wobei die $\Omega_{\varepsilon_\mu}(\tau)$ die den ε_μ entsprechenden Perioden des Integrals $\Phi(\tau)$ von $\varphi(\tau)$ sind. Also läßt sich $I(\varphi, \psi)$ einerseits als eine Linearform in den Koeffizienten der $\Omega_{\varepsilon_\mu}(\tau)$ für gewisse ε_μ schreiben. Andererseits ist aber $I(\varphi, \psi)$ auch wegen (26) eine Linearform in den Koeffizienten der Perioden Ω_{ε_μ} Integrals $\Psi(\tau)$ von $\psi(\tau)$. Das ergibt die erstere Behauptung.

$I(\varphi, \psi)$ läßt sich explizite berechnen, wenn keine Ecken und Spitzen vorhanden sind. Man geht dann von einer kanonischen Zerschneidung von $\mathfrak{R}$ mit der Seitenfolge $A_1, B_1, -\alpha_1 A_1, -\beta_1 B_1, A_2, \ldots, -\beta_g B_g$ aus; $\alpha_1, \beta_1, \ldots$ sollen die Bedeutung von § 1, Nr. 1 haben, und das Minuszeichen soll den der Integrationsrichtung entgegengesetzten Richtungssinn anzeigen. Man bezeichne die Ecke zwischen $-\beta_g B_g$ und A_1 mit τ_1, die Ecke zwischen A_1 und B_1 mit τ_2, und von dort ab fortlaufend mit τ_3, τ_4, usw. Alle diese Ecken entsprechen einem einzigen Punkt in $\mathfrak{R}$, d.h. die τ_μ gehen aus τ_1 durch Elemente der Gruppe Γ hervor. Und zwar findet man durch leichte Überlegung

$$\tau_2 = \alpha_1^{-1}\beta_1^{-1}\alpha_1\tau_1, \quad \tau_3 = \beta_1^{-1}\alpha_1\tau_1, \quad \tau_4 = \alpha_1\tau_1, \quad \tau_5 = \beta_1\alpha_1^{-1}\beta_1^{-1}\alpha_1\tau_1.$$

Nun ist in verständlicher Abkürzung

$$I(\varphi, \psi) = \sum_{\mu=1}^{g} I_\mu(\varphi, \psi) = \frac{1}{2\pi i} \sum_{\mu=1}^{g} \left(\int_{A_\mu} + \int_{B_\mu} - \int_{\alpha_\mu A_\mu} - \int_{\beta_\mu B_\mu} \right)$$

und

$$I_1(\varphi, \psi) = -\frac{1}{2\pi i} \int_{\tau_1}^{\alpha_1^{-1}\beta_1^{-1}\alpha_1\tau_1} \Omega_{\alpha_1}\, \psi\, d\tau - \frac{1}{2\pi i} \int_{\alpha_1^{-1}\beta_1^{-1}\alpha_1\tau_1}^{\beta_1^{-1}\alpha_1\tau_1} \Omega_{\beta_1}\, \psi\, d\tau.$$

Im zweiten Integral rechts substituiert man $\tau = \beta_1^{-1}\alpha_1\sigma$ und erhält dann unter Benutzung von (16)

$$I_1(\varphi,\psi) = -\frac{1}{2\pi i}\int_{\tau_1}^{\alpha_1^{-1}\beta_1^{-1}\alpha_1\tau_1} \Omega_{\alpha_1}\psi\,d\tau + \frac{1}{2\pi i}\int_{\tau_1}^{\alpha_1^{-1}\beta_1\alpha_1^{-1}\beta_1^{-1}\alpha_1\tau_1} (\Omega_{\alpha_1} - \Omega_{\beta_1^{-1}\alpha_1})\,\psi\,d\tau.$$

Die zweite der oben ausgesprochenen Behauptungen wird hiernach bewiesen sein, wenn man zeigt, daß das Integral

$$2\pi i\, I_{\eta,\varepsilon}(\varphi,\psi) = \int_{\tau_1}^{\varepsilon\tau_1} \Omega_\eta(\tau)\,\psi(\tau)\,d\tau$$

für zwei beliebige Elemente $\varepsilon, \eta \in \Gamma$ eine bilineare Form mit ganzen rationalen Koeffizienten in den Koeffizienten von $\Omega_\eta(\tau)$ und $\Omega'_{\varepsilon^{-1}}(\tau)$ ist. Dabei darf man die Integrationskonstanten von $\Psi(\tau)$ unterdrücken, da diese nach Nr. 1 in $I(\varphi,\psi)$ nicht eingehen; als untere Grenze des Integrals (12) nehme man $\tau_0 = \tau_1$. Es ist dann nach (14)

$$\Omega'_\varepsilon(\tau) = \frac{1}{(n-2)!}\int_{\tau_1}^{\varepsilon\tau_1} (\tau-\sigma)^{n-2}\psi(\sigma)\,d\sigma \cdot [\varepsilon]^{n-2}.$$

Daraus folgt nach (16)

$$\Omega'_{\varepsilon^{-1}}(\tau) = -\frac{1}{(n-2)!}\int_{\tau_1}^{\varepsilon\tau_1} (\tau-\sigma)^{n-2}\psi(\sigma)\,d\sigma.$$

Folglich ist

$$\Omega'_{\varepsilon^{-1}}(\tau) = \sum_{\nu=0}^{n-2} c'_\nu(\varepsilon^{-1})\,\tau^\nu = -\sum_{\nu=0}^{n-2} \tau^\nu \cdot \frac{\binom{n-2}{\nu}}{(n-2)!}\int_{\tau_1}^{\varepsilon\tau_1} (-\sigma)^{n-2-\nu}\psi(\sigma)\,d\sigma$$

oder

$$\int_{\tau_1}^{\varepsilon\tau_1} \sigma^\nu\psi(\sigma)\,d\sigma = -(-1)^\nu\nu!\,(n-2-\nu)!\,c'_{n-2-\nu}(\varepsilon^{-1}).$$

Setzt man endlich

$$\Omega_\eta(\tau) = \sum_{\nu=0}^{n-2} c_\nu(\eta)\,\tau^\nu,$$

so wird

$$2\pi i\, I_{\eta,\varepsilon}(\varphi,\psi) = -\sum_{\nu=0}^{n-2}(-1)^\nu\nu!\,(n-2-\nu)!\,c_\nu(\eta)\,c'_{n-2-\nu}(\varepsilon^{-1}),$$

womit alles bewiesen ist.

Die explizite Gestalt der Bilinearform $I(\varphi,\psi)$ sieht allerdings recht kompliziert aus. Einzig für $n=2$ erhält man ein einfaches Resultat, nämlich

$$I(\varphi,\psi) = \frac{1}{2\pi i}\sum_{\mu=1}^{g}\left(\Omega_{\alpha_\mu}\Omega'_{\beta_\mu} - \Omega_{\beta_\mu}\Omega'_{\alpha_\mu}\right) \qquad (n=2). \tag{27}$$

§ 4. Die automorphen Formen 2. Gattung

1. *Von jetzt ab wird bis zum Schluß n als gerade vorausgesetzt: $n = 2f$.* Das Hauptziel dieses Paragraphen ist eine Übersicht über sämtliche automorphen Formen 2. Gattung. Wir brauchen dazu den nachstehenden Hilfssatz, der aus einem allgemeineren Satze von A. WEIL [6], [7] gefolgert werden kann. Spezialfälle des Satzes von WEIL wurden übrigens später von O. TEICHMÜLLER ($n = 2$) und H. PETERSSON [5] (τ eine Grenzkreisuniformisierende) wiederentdeckt.

Es sei $\varphi(\tau)$ eine automorphe Form $(-n)$-ten Grades und 1. oder 2. Gattung. Für alle automorphen Formen 1. Gattung $\psi(\tau)$ dieses Grades gelte $I(\varphi, \psi) = 0$. Dann ist $\varphi(\tau) = \varphi_0(\tau) + \varphi_1(\tau)$ mit einer automorphen Form $\varphi_0(\tau)$ von 1. Gattung und der $(n-1)$-fachen Ableitung $\varphi_1(\tau)$ einer automorphen Form $\Phi_1(\tau)$ vom Grade $n-2$.

Wir bringen einen vollständigen Beweis, da wir auf ihn weiter unten Bezug nehmen müssen. Sei $\Phi(\tau)$ das Integral (12) von $\varphi(\tau)$. Es wird eine automorphe Form $\Phi_1(\tau)$ des Grades $n-2$ gesucht, derart daß

$$\Phi_0(\tau) = \Phi(\tau) - \Phi_1(\tau) \tag{28}$$

höchstens an einer (von den Ecken $\mathfrak{e}_\lambda$ und Spitzen $\mathfrak{s}_\nu$ verschiedenen) Stelle singulär ist.

Zu einer solchen Form $\Phi_1(\tau)$ gehört ein Divisor $\mathfrak{F}$, der durch einen gegebenen Primdivisor $\mathfrak{p}$ ebensooft teilbar ist wie $\Phi_1(\tau)\left(\frac{dz}{d\tau}\right)^{f-1}$, wo z eine lokale Uniformisierende zu $\mathfrak{p}$ ist; $\mathfrak{F}$ gehört der $(1-f)$-ten Potenz der kanonischen Klasse W an. Nach § 1, Nr. 1 ist $\frac{dz}{d\tau}$ durch ein gewöhnliches $\mathfrak{p}$ gar nicht teilbar, durch die $\mathfrak{s}_\nu$ genau einmal, und durch die $\mathfrak{e}_\lambda$ genau $\left(1 - \frac{1}{n_\lambda}\right)$-mal. Hat nun $\Phi_1(\tau)$ Pole der Ordnungen m_μ in den Punkten $\mathfrak{q}_\mu$, so ist der Divisor $\mathfrak{F}$:

$$\mathfrak{F} = \mathfrak{G} \cdot \prod_\mu \mathfrak{q}_\mu^{-m_\mu} \cdot \prod_{\nu=1}^{h} \mathfrak{s}_\nu^{f-1} \cdot \prod_{\lambda=1}^{k} \mathfrak{e}_\lambda^{-\left[\left(\frac{1}{n_\lambda} - 1\right)(f-1)\right]}$$

mit einem ganzen Divisor $\mathfrak{G}$. Da $\mathfrak{F}$ zur Klasse W^{1-f} gehört, ist der Grad von $\mathfrak{G}$

$$\sum_\mu m_\mu - h(f-1) - \sum_{\lambda=1}^{k}\left(f - 1 - \left[\frac{f-1}{n_\lambda}\right]\right) - 2(f-1)(g-1)$$

und die Dimension der Klasse von $\mathfrak{G}$ nach dem Riemann-Rochschen Satze

$$\dim \mathfrak{G} = \sum_\mu m_\mu - (2f-1)(g-1) - h(f-1) - \sum_{\lambda=1}^{k}\left(f - 1 - \left[\frac{f-1}{n_\lambda}\right]\right) + \dim \frac{W}{\mathfrak{G}},$$

wo W die kanonische Klasse bezeichnet.

Diese Dimension ist >0, wenn $\sum m_\mu > G_n$, wie ein Vergleich mit der Formel (11) lehrt. Unter dieser Voraussetzung gibt es also eine automorphe

Form $\Phi_1(\tau)$ $(n-2)$-ten Grades, die überall holomorph bis auf Pole höchstens m_μ-ter Ordnung in den Punkten $\mathfrak{q}_\mu$ ist.

Wir müssen speziell die Existenz einer Form $\Phi_1(\tau)$ nachweisen, die in einem gegebenen Punkt $\mathfrak{q}$ einen Pol genau m-ter Ordnung und in einem weiteren Punkt $\mathfrak{q}_0$, über den noch zu verfügen ist, einen Pol höchstens G_n-ter Ordnung hat. Damit $\Phi_1(\tau)$ diese Eigenschaft hat, muß der Divisor $\mathfrak{G}$ zu $\mathfrak{q}$ teilerfremd sein. Einen solchen Divisor $\mathfrak{G}$ gibt es dann und nur dann, wenn die Dimension der Klasse von $\mathfrak{G}\mathfrak{q}^{-1}$ um 1 kleiner ist als die von $\mathfrak{G}$. Nach dem Riemann-Rochschen Satze ist unter den gemachten Voraussetzungen

$$\dim \mathfrak{G} = m + \dim W\mathfrak{G}^{-1}, \qquad \dim \mathfrak{G}\mathfrak{q}^{-1} = m - 1 + \dim W\mathfrak{G}^{-1}\mathfrak{q}.$$

Die Klasse von $W\mathfrak{G}^{-1}$ ist die von $W'\mathfrak{P}^{-1}\mathfrak{q}_0^{-G_n}\mathfrak{q}^{-m}$:

$$W\mathfrak{G}^{-1} \sim W'\mathfrak{P}\mathfrak{q}_0^{-G_n}\mathfrak{q}^{-m}, \qquad \mathfrak{P} = \prod_{\nu=1}^{n}\mathfrak{p}_\nu^{l-1}\prod_{\lambda=1}^{k}\mathfrak{e}_\lambda^{-\left[\left(\frac{1}{n_\lambda}-1\right)(l-1)\right]}.$$

Wegen (10) hat $\mathfrak{P}$ dieselbe Bedeutung wie in § 1, Nr. 3.

Wir behaupten nun, daß für fast alle $\mathfrak{q}_0$ die Dimension der Klasse $W\mathfrak{G}^{-1}\mathfrak{q}^m$ und daher erst recht die Dimensionen der Klassen $W\mathfrak{G}^{-1}\mathfrak{q}$ und $W\mathfrak{G}^{-1}$ gleich 0 sind, woraus $\dim \mathfrak{G} = \dim \mathfrak{G}\mathfrak{q}^{-1} + 1$ und die Existenz der Form $\Phi_1(\tau)$ folgen würde.

Ist $\dim W\mathfrak{G}^{-1}\mathfrak{q}^m > 0$, so gibt es einen ganzen Divisor $\mathfrak{h} \sim W'\mathfrak{P}\mathfrak{q}_0^{-G_n}$. Dann ist $\mathfrak{h}\mathfrak{q}_0^{G_n}$ ein in der Klasse $W'\mathfrak{P}$ gelegener ganzer und durch $\mathfrak{q}_0^{G_n}$ teilbarer Divisor. Zu ihm gehört nach § 1, Nr. 3 eine überall holomorphe automorphe Spitzenform $\chi_1(\tau)$ des Grades $-n$, die in $\mathfrak{q}_0$ von der Ordnung G_n verschwindet. Wir ergänzen nun $\chi_1(\tau)$ zu einer Basis $\chi_i(\tau)$ des Moduls aller Spitzenformen dieses Grades und bilden die Determinante

$$D(\tau) = \begin{vmatrix} \chi_1(\tau) & \frac{d}{d\tau}\chi_1(\tau) & \cdots & \left(\frac{d}{d\tau}\right)^{G_n-1}\chi_1(\tau) \\ \cdots & \cdots & \cdots & \cdots \\ \chi_{G_n}(\tau) & \frac{d}{d\tau}\chi_{G_n}(\tau) & \cdots & \left(\frac{d}{d\tau}\right)^{G_n-1}\chi_{G_n}(\tau) \end{vmatrix}.$$

Dieses ist offenbar eine nicht identisch verschwindende automorphe Form des Grades $-\frac{G_n(G_n+2n-1)}{2}$. Da $\chi_1(\tau)$ in $\mathfrak{q}_0$ von der Ordnung G_n verschwinden sollte, ist auch $D(\tau) = 0$ in $\mathfrak{q}_0$. Das ist aber nur für endlich viele Punkte $\mathfrak{q}_0$ möglich.

Wir halten noch ein Nebenergebnis fest. Ist $D(\tau) \neq 0$ in $\mathfrak{q}_0$, und gelten in einem $\mathfrak{q}_0$ entsprechenden Punkte τ_0 der τ-Ebene die Reihenentwicklungen

$$\chi_i(\tau) = \sum_{\nu=0}^{\infty} c_{i\nu}(\tau - \tau_0)^\nu,$$

so ist die Determinante

$$|c_{i\nu}| = \frac{1}{1!\,2!\ldots(G_n-1)!}D(\tau_0) \qquad (i = 1, \ldots, G_n;\ \nu = 0, 1, \ldots, G_n - 1).$$

Mit der zu $(c_{i\nu})$ reziproken Matrix $(c'_{i\nu})$ setze man $\varphi_i(\tau) = \sum_\nu c'_{i\nu}\chi_\nu(\tau)$. Diese Funktionen bilden dann eine Basis des Moduls der holomorphen Spitzenformen mit der Eigenschaft, daß $\varphi_i(\tau)$ in $\mathfrak{q}_0$ genau von der Ordnung $i-1$ verschwindet, $i=1, \ldots, G_n$. Von dieser Basis wird noch Gebrauch gemacht werden.

Nach diesen Vorbereitungen kommen wir zu dem eigentlichen Beweis des Hilfssatzes. Das Integral $\Phi(\tau)$ habe Pole m_μ-ter Ordnung in $\mathfrak{q}_\mu$ $(\mu=1, 2, \ldots)$. Man nehme einen weiteren Punkt $\mathfrak{q}_0 \neq \mathfrak{q}_\mu, \mathfrak{s}_\nu, \mathfrak{e}_\lambda$ hinzu und konstruiere automorphe Formen $\Phi_{\mu i}(\tau)$ $(\mu=1, 2, \ldots;\ i=1, 2, \ldots, m_\mu)$ mit einem Pol der Ordnung i in $\mathfrak{q}_\mu$ und einem Pol einer Ordnung $\leq G_n$ in $\mathfrak{q}_0$. Eine passende Linearkombination $\Phi_1(\tau)$ der $\Phi_{\mu i}(\tau)$ erlaubt dann, die Pole von $\Phi(\tau)$ zu kompensieren, derart, daß die Differenz (28) überall bis möglicherweise auf die Stelle $\mathfrak{q}_0$ holomorph ist, wo ein Pol der Ordnung $\leq G_n$ liegen darf.

Nun wird von der Voraussetzung $I(\varphi, \psi)=0$ für alle automorphen Formen $\psi(\tau)$ 1. Gattung Gebrauch gemacht. Für $\psi(\tau)$ werden die soeben konstruierten Formen $\varphi_i(\tau)$ eingesetzt. Aus der Voraussetzung folgt wegen § 3, Nr. 2:

$$\int_{\mathfrak{R}} \Phi_0(\tau)\,\varphi_i(\tau)\,d\tau = 0.$$

Der Integrand hat als einzige Singularität einen Pol der Ordnung $\leq G_n - i + 1$ in $\mathfrak{q}_0$, und dessen Residuum ist 0. Da dieses für alle $i=1, 2, \ldots, G_n$ gilt, muß $\Phi_0(\tau)$ auch in $\mathfrak{q}_0$ holomorph sein, und der Hilfssatz ist bewiesen.

2. *Es sei n gerade. Zu einer Basis $\varphi_i(\tau)$ $(i=1, \ldots, G_n)$ des Moduls der automorphen Formen $(-n)$-ten Grades und 1. Gattung gibt es ebenso viele automorphe Formen $\psi_i(\tau)$ 2. Gattung derart, daß*

$$I(\varphi_i, \psi_k) = \begin{cases} 1 \text{ für } i=k, \\ 0 \text{ für } i \neq k, \end{cases} \qquad I(\psi_i, \psi_k) = 0 \tag{29}$$

ist. Jede automorphe Form $\varphi(\tau)$ 1. oder 2. Gattung läßt sich in eindeutiger Weise als Linearkombination

$$\varphi(\tau) = \sum_{i=1}^{G_n} p_i\varphi_i(\tau) + \sum_{i=1}^{G_n} q_i\psi_i(\tau) + \left(\frac{d}{d\tau}\right)^{n-1} \Phi_0(\tau) \tag{30}$$

darstellen, wobei $\Phi_0(\tau)$ eine automorphe Form des Grades $n-2$ ist.

Es genügt offenbar, diese Behauptung für eine bestimmte Basis $\varphi_i(\tau)$ zu beweisen. Wir nehmen eine solche, für die $\varphi_i(\tau)$ an einer Stelle $\mathfrak{q}_0 \neq \mathfrak{s}_\nu, \mathfrak{e}_\lambda$ genau von der Ordnung $i-1$ verschwindet (s. Nr. 1). Unter Benutzung des Riemann-Rochschen Satzes kann man leicht eine automorphe Spitzenform $\chi_N(\tau)$ konstruieren (s. dazu § 1, Nr. 3), welche überall holomorph bis auf einen Pol N-ter Ordnung in $\mathfrak{q}_0$ ist. Aus $\chi_1(\tau), \ldots, \chi_N(\tau)$ läßt sich eine automorphe Spitzenform $\psi'_N(\tau)$ zusammensetzen, deren Pol in einer $\mathfrak{q}_0$ entsprechenden Stelle τ_0 den Hauptteil $(\tau-\tau_0)^{-N}$ hat. Für $N=n, n+1, \ldots$ sind die $\psi'_N(\tau)$ von 2. Gattung. Setzt man nun $\psi_k(\tau) = \psi'_{n+k}(\tau)$ $(k=1, \ldots, G_n)$, so ist die Determinante $I(\varphi_i, \psi_k) \neq 0$.

Durch geeignete lineare Transformation der $\psi_k(\tau)$ kann man jetzt $I(\varphi_i, \psi_k) = 1$ bzw. 0 erreichen. Wegen § 3, Nr. 2 bleiben diese Gleichungen ungeändert, wenn man die ψ_k durch $\psi_k + \varphi_k'$ mit beliebigen automorphen Formen $\varphi_k' = \sum_i c_{ki} \varphi_i$ ersetzt. Es ist

$$I(\psi_i + \varphi_i', \psi_k + \varphi_k') = I(\psi_i, \psi_k) + I(\psi_i, \varphi_k') + I(\varphi_i', \psi_k) = I(\psi_i, \psi_k) + c_{ik} - c_{ki}.$$

Wegen (26) lassen sich die c_{ik} so bestimmen, daß dies gleich 0 wird. Damit ist der erste Teil der Behauptung bewiesen.

Jetzt sei $\varphi(\tau)$ eine beliebige automorphe Form 1. oder 2. Gattung. Man bilde

$$\varphi_0'(\tau) = \varphi(\tau) + \sum_{i=1}^{G_n} I(\varphi, \varphi_i)\, \psi_i(\tau).$$

Wegen (29) ist $I(\varphi_0', \varphi_i) = 0$ für alle i und daher nach dem Hilfssatz in Nr. 1 $\varphi_0'(\tau)$ die Summe einer automorphen Form $\sum_i p_i \varphi_i(\tau)$ 1. Gattung und der Ableitung einer automorphen Form $\Phi_0(\tau)$ des Grades $n-2$.

3. ***Für ein gerades n entsprechen sich die Cohomologieklassen der Perioden Ω_ε und die Klassen automorpher Formen 1. oder 2. Gattung modulo den $(n-1)$-fachen Ableitungen automorpher Formen vom Grade $n-2$ umkehrbar eindeutig. (Oder auch die Cohomologieklassen und die Klassen der Integrale 1. oder 2. Gattung modulo den automorphen Formen des Grades $n-2$.)***

Dieses geht durch Vergleich der Anzahlen (11) und (24) hervor. Nach dem Satz in Nr. 2 ist nämlich die Anzahl der Klassen der automorphen Formen 1. oder 2. Gattung modulo den $(n-1)$-fachen Ableitungen gleich $2G_n$, und diese Zahl ist nach (11) und (24) gleich der Anzahl der linear unabhängigen Cohomologieklassen.

Allerdings ist (24) noch nicht allgemein bewiesen worden, und wir müssen den Rest des Beweises nachholen. Es handelt sich um den Fall $g = h = 0$. Die Potenzprodukte $\varepsilon_1^{h_{11}} \varepsilon_2^{h_{12}} \dots \varepsilon_1^{h_{21}} \varepsilon_2^{h_{22}} \dots$ mit $h_{1\lambda} + h_{2\lambda} + \dots \equiv 0 \bmod n_\lambda$ $(\lambda = 1, \dots, k)$ bilden eine Untergruppe Γ_0 von Γ von endlichem Index m. Γ_0 enthält kein Element von endlicher Ordnung mehr, da alle Elemente von endlicher Ordnung in Γ die Konjugierten der ε_λ und deren Potenzen sind. Für Γ_0 an Stelle von Γ ist obige Behauptung also richtig. Eine vorgelegte Cohomologieklasse für Γ ist natürlich gleichzeitig eine für Γ_0, und zu dieser existiert also ein Integral $\Phi_0(\tau)$ von 1. oder 2. Gattung, welches den Gln. (13) für alle $\varepsilon_0 \in \Gamma_0$ mit den gegebenen Perioden Ω_{ε_0} genügt.

Es sei nun $\Gamma = \sum_{i=1}^{m} \Gamma_0 \varepsilon_i$ eine Zerlegung in Nebengruppen und

$$\Phi(\tau) = \sum_{i=1}^{m} \Phi_0(\tau)\, [\varepsilon_i]^{n-2}.$$

Für jedes $\varepsilon \in \Gamma$ gibt es dann ein $\varepsilon_0(\varepsilon, i) \in \Gamma_0$ und einen Index $1 \leq j(\varepsilon, i) \leq m$ so daß

$$\varepsilon_i \varepsilon = \varepsilon_0(\varepsilon, i)\, \varepsilon_{j(\varepsilon, i)} \tag{31}$$

ist. Nun folgt

$$\Phi(\tau)[\varepsilon]^{n-2} = \sum_{i=1}^{m} \Phi_0(\tau)[\varepsilon_0(\varepsilon,i)]^{n-2}[\varepsilon_{j(\varepsilon,i)}]^{n-2}$$

$$= \sum_{i=1}^{m} \Phi_0(\tau)[\varepsilon_{j(\varepsilon,i)}]^{n-2} + \sum_{i=1}^{m} \Omega_{\varepsilon_0(\varepsilon,i)}(\tau)[\varepsilon_{j(\varepsilon,i)}]^{n-2}.$$

Offenbar durchläuft $j(\varepsilon, i)$ mit i auch alle Ziffern von 1 bis m. Daher ist

$$\Phi(\tau)[\varepsilon]^{n-2} = \Phi(\tau) + \sum_{j=1}^{m} \Omega_{\varepsilon_0(\varepsilon,i)}(\tau)[\varepsilon_{j(\varepsilon,i)}]^{n-2}$$

und weiter nach (16)

$$\frac{1}{m}\Phi(\tau)[\varepsilon]^{n-2} = \frac{1}{m}\Phi(\tau) + \Omega_\varepsilon(\tau) + \Big(\sum_{i=1}^{m}\Omega_{\varepsilon_i}(\tau)\Big)\big([\varepsilon]^{n-2} - [1]^{n-2}\big)$$

$$= \frac{1}{m}\Phi(\tau) + \Omega'_\varepsilon(\tau).$$

$\Omega'_\varepsilon(\tau)$ und $\Omega_\varepsilon(\tau)$ gehören der gleichen Cohomologieklasse an, und $\frac{1}{m}\Phi(\tau)$ ist ein zu dieser gehöriges Integral 1. oder 2. Gattung.

Hiermit ist obiger Satz auch in diesem Falle bewiesen, und die Formel (24) kann man nachträglich aus (11) gewinnen, da man jetzt weiß, daß die Anzahl der linear unabhängigen Cohomologieklassen $R_n = 2G_n$ ist.

§ 5. Korrespondenzen

1. In diesem Paragraphen wird der Begriff und die elementaren Eigenschaften gewisser Korrespondenzen entwickelt (es handelt sich um spezielle irreduzible Korrespondenzen), und zwar vom funktionentheoretischen oder auch vom geometrischen Standpunkt. Die algebraischen Gesichtspunkte der Korrespondenzentheorie interessieren in diesem Zusammenhang weniger.

Es möge eine zweireihige nicht singuläre und nicht zu Γ gehörige Matrix ν_0 existieren, für die $\Gamma \cap \nu_0^{-1}\Gamma\nu_0$ sowohl in Γ wie in $\nu_0^{-1}\Gamma\nu_0$ einen endlichen Index hat.

$$\Gamma = \sum_{i=1}^{d} (\Gamma \cap \nu_0^{-1}\Gamma\nu_0)\,\varepsilon_i, \qquad \nu_0^{-1}\Gamma\nu_0 = \sum_{i=1}^{d^*} (\Gamma \cap \nu_0^{-1}\Gamma\nu_0)\,\varepsilon_i^* \tag{32}$$

seien die Entwicklungen in linksseitige Nebengruppen. Zu jedem $\varepsilon \in \Gamma$ und zu jeder Ziffer i zwischen 1 und d gibt es ein Element $\eta(\varepsilon, i) \in \Gamma \cap \nu_0^{-1}\Gamma\nu_0$ und eine Ziffer $j(\varepsilon, i)$ zwischen 1 und d, so daß

$$\varepsilon_i \cdot \varepsilon = \eta(\varepsilon, i)\,\varepsilon_{j(\varepsilon,i)}, \qquad \eta(\varepsilon, i) \in \Gamma \cap \nu_0^{-1}\Gamma\nu_0 \tag{33}$$

gilt; die Abbildung $i \to j(\varepsilon, i)$ ist eine Permutation der Ziffern 1 bis d. Man bilde noch die Nebengruppen

$$\Gamma\nu_i \quad \text{mit} \quad \nu_i = \nu_0\varepsilon_i, \qquad i = 1, \ldots, d. \tag{34}$$

Wegen (33) permutieren sie sich, wenn man sie von rechts mit einem $\varepsilon \in \Gamma$ multipliziert.

Die Beziehungen zwischen den Gruppen $\Gamma, \Gamma \cap \nu_0^{-1}\Gamma\nu_0, \nu_0^{-1}\Gamma\nu_0$ lassen eine bedeutungsvolle geometrische Deutung zu. Nach § 1, Nr. 1 ist Γ eine eigentlich diskontinuierliche Gruppe von Abbildungen einer einfach zusammenhängenden Riemannschen Fläche $\mathfrak{T}$ auf sich. Man bilde die Klassen $\Gamma\tau$ von bezüglich Γ äquivalenten Punkten von $\mathfrak{T}$. Diese bilden eine Fläche, welche als der Quotientenraum $\Gamma\backslash\mathfrak{T}$ gedeutet werden kann. Diese Fläche ist mit $\mathfrak{R}$ homöomorph, ja sogar von den Spitzen und Ecken abgesehen mit $\mathfrak{R}$ konform äquivalent. Wir dürfen daher

$$\Gamma\backslash\mathfrak{T} = \mathfrak{R}$$

schreiben.

Der Quotientenraum $(\Gamma \cap \nu_0^{-1}\Gamma\nu_0)\backslash\mathfrak{T} = \overline{\mathfrak{R}}$ ist eine d-blättrige Überlagerung von $\mathfrak{R}$, wenn man die Punkte $\overline{P}_i = (\Gamma \cap \nu_0^{-1}\Gamma\nu_0)\,\varepsilon_i\tau$ von $\overline{\mathfrak{R}}$ als diejenigen erklärt, welche über dem Punkt $P = \Gamma\tau$ von $\mathfrak{R}$ liegen. In gleicher Weise liegen die $\overline{P}_i$ über den Punkten $P_i^* = \nu_0^{-1}\Gamma\nu_0 \cdot \overline{P}_i = \nu_0^{-1}\Gamma\nu_0\varepsilon_i\tau$ der Fläche $\nu_0^{-1}\Gamma\nu_0\backslash\mathfrak{T} = \mathfrak{R}^*$. Endlich wird durch $\nu_0 P^* = \nu_0(\nu_0^{-1}\Gamma\nu_0\tau) = \Gamma\nu_0\tau = P$ jeder Punkt P^* von $\mathfrak{R}^*$ auf einen Punkt P von $\mathfrak{R}$ abgebildet. Die zusammengesetzte Abbildung

$$P \to \{\overline{P}_1, \ldots, \overline{P}_d\} \to \{P_1^*, \ldots, P_d^*\} \to \{P_1, \ldots, P_d\}$$

eines Punktes P von $\mathfrak{R}$ auf eine Punktgruppe P_i von $\mathfrak{R}$ nennt man eine *Korrespondenz von $\mathfrak{R}$ zu sich selber*. Die Abbildungen $P \to P_i$ sind natürlich fast überall konform, ausgenommen sind höchstens die Stellen, wo $\overline{\mathfrak{R}}$ über $\mathfrak{R}$ oder $\mathfrak{R}^*$ verzweigt ist.

Die soeben erklärte Korrespondenz läßt sich als ein linksseitiger Operator schreiben:

$$T(\Gamma\tau) = \Big(\sum_{i=1}^{d} \Gamma\nu_i\Big)(\Gamma\tau). \tag{35}$$

In der Tat werden jedem Punkt $\varepsilon\tau$ von $\mathfrak{T}$ die Punkte $(\Gamma \cap \nu^{-1}\Gamma\nu)\,\varepsilon_i\,\varepsilon\tau$ von $\overline{\mathfrak{R}}$, die Punkte $\nu_0^{-1}\Gamma\nu_0 \cdot (\Gamma \cap \nu_0^{-1}\Gamma\nu_0)\,\varepsilon_i\,\varepsilon\tau = \nu_0^{-1}\Gamma\nu_0\varepsilon_i\,\varepsilon\tau$ von $\mathfrak{R}^*$ und die Punkte $\nu_0 \cdot \nu_0^{-1}\Gamma\nu_0\varepsilon_i\,\varepsilon\tau = \Gamma\nu_0\varepsilon_i\,\varepsilon\tau = \Gamma\nu_i\varepsilon\tau$ von $\mathfrak{R}$ zugeordnet. Diese sind von ε unabhängig, so daß man (35) in dem Sinne schreiben kann, daß die Summe als mengentheoretische Vereinigung aufzufassen ist.

Die Schreibweise (35) legt die Definition der *Summe* und des *Produktes von zwei Korrespondenzen* nahe. Es entstehen so weitere Abbildungen eines Punktes von $\mathfrak{R}$ auf Punktgruppen von $\mathfrak{R}$, die allerdings nicht mehr in obiger Weise durch Gruppen $\nu_0^{-1}\Gamma\nu_0$ erzeugbar zu sein brauchen. Dieses sind Korrespondenzen allgemeinerer Art, die hier nicht weiter interessieren, und auf die wir deshalb auch nicht näher eingehen werden. Es ist jedenfalls klar, daß diese allgemeineren Korrespondenzen einen assoziativen Ring bilden.

2. Man kann die Reihenfolge $\mathfrak{R} \to \overline{\mathfrak{R}} \to \mathfrak{R}^* \to \mathfrak{R}$ der Abbildungen auch umkehren. Dann bildet man zuerst einen Punkt $P = \Gamma\tau$ von $\mathfrak{R}$ durch $\nu_0^{-1}P = \nu_0^{-1}\Gamma\tau = \nu_0^{-1}\Gamma\nu_0 \cdot \nu_0^{-1}\tau = P^*$ auf einen Punkt P^* von $\mathfrak{R}^*$ ab. Hierauf sucht man dessen Überlagerungen $\overline{P}_i = (\Gamma \cap \nu_0^{-1}\Gamma\nu_0)\,\varepsilon_i^*\nu_0^{-1}\tau$ $(i = 1, \ldots, d^*)$ in $\overline{\mathfrak{R}}$ und schließlich diejenigen Punkte $P_i = \Gamma\overline{P}_i = \Gamma\varepsilon_i^*\nu_0^{-1}\tau$ in $\mathfrak{R}$ auf, über welchen

die $\overline{P}_i$ liegen. Diese Punktgruppe bezeichnet man mit

$$T^*(\Gamma\tau) = \sum_{i=1}^{d^*} \Gamma \varepsilon_i^* \nu_0^{-1} \tau.$$

Transformiert man die zweite Gl. (32) noch mit ν_0, so entsteht

$$\Gamma = \sum_{j=1}^{d^*} (\Gamma \cap \nu_0 \Gamma \nu_0^{-1}) \eta_j, \tag{36}$$

und T läßt sich schreiben

$$T^*(\Gamma\tau) = \left(\sum_{j=1}^{d^*} \Gamma \nu_j^* \right) (\Gamma\tau) \quad \text{mit } \nu_j^* = \nu_0^{-1} \eta_j. \tag{37}$$

T^* heißt die zu T *adjungierte Korrespondenz*. Offenbar ist

$$T^{**} = T.$$

Es läßt sich sehr leicht zeigen, *daß die Konjugiertenbildung ein Antiautomorphismus im Ring der Korrespondenzen ist.* Wir werden aber hiervon keinen Gebrauch machen und auch den Beweis nicht bringen.

3. *Ein Beispiel* soll das bisher Ausgeführte verdeutlichen. Es sei Γ die Modulgruppe. Dann ist $\mathfrak{T}$ die obere komplexe Halbebene. Spezielle Korrespondenzen sind die *Modularkorrespondenzen* $T = T_p$; sie lassen sich jeder natürlichen Primzahl p zuordnen. Man nehme

$$\nu_0 = \nu_1 = \begin{pmatrix} p & 0 \\ 0 & 1 \end{pmatrix}, \quad \nu_2 = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}^{-1} \nu_0 \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix},$$

$$\nu_{2+i} = \nu_2 \begin{pmatrix} 1 & i \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & i \\ 0 & p \end{pmatrix} \quad (i = 0, \ldots, p-1).$$

Hier sind die ν_i auch noch auf der linken Seite mit Elementen aus Γ multipliziert worden, was aber ersichtlich auf die Definition (35) keinen Einfluß hat. Modularkorrespondenzen T_N mit quadratfreiem N sind von ähnlicher Natur. Bei nicht quadratfreiem N sind sie bereits additiv aus Korrespondenzen obiger Art zusammengesetzt.

4. Die Korrespondenzen lassen sich auf folgende Art als Operatoren für den Modul der automorphen Formen eines beliebigen Grades m erklären:

$$\varphi(\tau)\, T = \sum_{i=1}^{d} \varphi(\tau)\, [\nu_i]^m. \tag{38}$$

Die rechte Seite ändert sich nicht, wenn man die ν_i links mit irgendwelchen Elementen $\eta \in \Gamma$ multipliziert. Wendet man aber auf $\varphi(\tau)\, T$ den Operator $[\varepsilon]^m$ an, so erkennt man die Invarianz auf Grund von (33). Speziell interessieren hier die Fälle $m = -n$ und $m = n - 2$.

Offensichtlich führt der durch (38) erklärte Operator holomorphe Formen in holomorphe und meromorphe Formen in ebensolche über. Wir behaupten sogar, daß Spitzenformen in Spitzenformen transformiert werden.

Hierzu ist zuerst zu zeigen, daß bei einer Korrespondenz T die Bildpunkte einer Spitze sämtlich wieder Spitzen sind. Nach § 1, Nr. 1 sind die

Spitzen dadurch eindeutig gekennzeichnet, daß sie Fixpunkte solcher Substitutionen $\gamma \in \Gamma$ sind, welche sich auf die Normalform (3) transformieren lassen. Ist $P = \Gamma\tau_0$ eine Spitze und $\gamma\tau_0 = \tau_0$, so sind die Bildpunkte $\overline{P}_i = (\Gamma \cap \nu_0^{-1}\Gamma\nu_0)\,\varepsilon_i\tau_0$. Alle Punkte $\eta\varepsilon_i\tau_0$ mit $\eta \in \Gamma \cap \nu_0^{-1}\Gamma\nu_0$ sind Fixpunkte der Substitutionen $\bar{\gamma} = \eta\varepsilon_i\gamma^l\varepsilon_i^{-1}\eta^{-1}$, wobei die Exponenten l so gewählt werden müssen, daß $\bar{\gamma} \in \Gamma \cap \nu_0^{-1}\Gamma\nu_0$ ist. Wegen der Endlichkeit des Index $[\Gamma : \Gamma \cap \nu_0^{-1}\Gamma\nu_0]$ gibt es solche Exponenten l. Die $\bar{\gamma}$ lassen sich auf die Normalform (3) transformieren, und daher sind die Punkte $\overline{P}_i$ Spitzen von $\mathfrak{R}$. Die Punkte $P_i^* = \nu_0^{-1}\Gamma\nu_0\overline{P}_i$ sind wegen dieser Eigenschaft auch Spitzen von $\mathfrak{R}^*$. Endlich bildet $\nu_0 P^* = P$ gleichzeitig einen Fundamentalbereich $\mathfrak{F}^*$ von $\nu_0^{-1}\Gamma\nu_0$ in $\mathfrak{T}$ auf einen Fundamentalbereich $\mathfrak{F}$ von Γ in $\mathfrak{R}$ überall konform ab; dabei müssen Spitzen in Spitzen übergehen.

Auf Grund von (38) kann man leicht schließen, daß eine in allen Spitzen verschwindende Funktion $\varphi(\tau)$ diese Eigenschaft bei Anwendung des Operators T behält. Es gilt daher:

Der Operator (38) *transformiert automorphe Formen 1. und 2. Gattung in automorphe Formen der gleichen Gattung.*

Da die automorphen Formen der 1. Gattung einen Modul endlichen Ranges bilden, erhält man so eine Darstellung des Korrespondenzenringes durch Matrizen eines endlichen Ranges. Wir stellen uns die Aufgabe, deren Spur zu berechnen; sie wird in § 6 gelöst.

5. Jetzt soll das Verhalten des Operators (38) bei $(n-1)$-facher unbestimmter Integration studiert werden. In der Bezeichnungsweise von § 1 ist

$$\Phi'(\tau) = \frac{1}{(n-2)!}\int_{\tau_0}^{\tau}(\tau-\sigma)^{n-2}\big(\varphi(\sigma)\,T\big)\,d\sigma$$

$$= \frac{1}{(n-2)!}\sum_{i=1}^{d}\int_{\tau_0}^{\tau}(\tau-\sigma)^{n-2}\varphi(\nu_i\sigma)\left(\frac{a_{\nu_i}d_{\nu_i}-b_{\nu_i}c_{\nu_i}}{(c_{\nu_i}\tau+d_{\nu_i})^2}\right)^{\frac{n-2}{2}}d\nu_i\sigma.$$

Man benutzt hier die Identität

$$\nu_i\tau - \nu_i\sigma = (\tau-\sigma)\,\frac{a_{\nu_i}d_{\nu_i}-b_{\nu_i}c_{\nu_i}}{(c_{\nu_i}\tau+d_{\nu_i})\,(c_{\nu_i}\sigma+d_{\nu_i})}$$

und erhält dann

$$\Phi'(\tau) = \frac{1}{(n-2)!}\sum_{i=1}^{d}\int_{\tau_0}^{\tau}(\nu_i\tau-\nu_i\sigma)^{n-2}\varphi(\nu_i\sigma)\,d\nu_i\sigma\left(\frac{(c_{\nu_i}\tau+d_{\nu_i})^2}{a_{\nu_i}d_{\nu_i}-b_{\nu_i}c_{\nu_i}}\right)^{\frac{n-2}{2}}$$

Substituiert man in den einzelnen Integralen $\nu_i\sigma = \sigma_i$, schreibt aber wieder σ an Stelle von σ_i, so wird dies

$$\Phi'(\tau) = \sum_{i=1}^{d}\Phi(\tau)\,[\nu_i]^{n-2} + \frac{1}{(n-2)!}\sum_{i=1}^{d}\int_{\nu_i\tau_0}^{\tau_0}\left(\frac{(a_{\nu_i}\tau+b_{\nu_i})-\sigma(c_{\nu_i}\tau+d_{\nu_i})^2}{a_{\nu_i}d_{\nu_i}-b_{\nu_i}c_{\nu_i}}\right)^{\frac{n-2}{2}}\varphi(\sigma)\,d\sigma$$

$$= \sum_{i=1}^{d}\Phi(\tau)\,[\nu_i]^{n-2} + \Theta(\tau).$$

Definiert man also T als einen Operator im Modul der Integrale 1. oder 2. Gattung durch

$$\Phi(\tau)\,T = \sum_{i=1}^{d} \Phi(\tau)\,[v_i]^{n-2}, \tag{39}$$

so erweisen sich die Anwendung von T und die $(n-1)$-fache Integration bzw. Differentiation als vertauschbare Operatoren, abgesehen von den Integrationskonstanten.

Natürlich ist mit (39) T auch als ein Operator im Raume der Cohomologieklassen erklärt. Die Perioden von $\Phi(\tau)\,T$ sind

$$\Omega_\varepsilon T = \Omega'_\varepsilon = \sum_{i=1}^{d} \Omega_{\eta(\varepsilon,i)}\,[v_{j(\varepsilon,i)}]^{n-2}, \tag{40}$$

wo $\eta(\varepsilon, i), j(\varepsilon, i), v_i$ durch (33) und (34) erklärt sind. Auf eine merkwürdige Konsequenz dieses Sachverhalts werden wir in § 7 hinweisen.

6. Zum Schluß dieses Paragraphen soll das Verhalten der Bilinearform $I(\varphi, \psi)$ bei Anwendung der Korrespondenz T studiert werden. Das Ergebnis sei an den Anfang gestellt: *Für zwei automorphe Formen 1. oder 2. Gattung gilt*

$$I(\varphi, \psi\,T) = I(\varphi\,T^*, \psi), \tag{41}$$

wo T^* die zu T konjugierte Korrespondenz ist.

Zum Beweis deute man die irgendwie aufgeschnittene Fläche $\mathfrak{R}$ als einen Fundamentalbereich $\mathfrak{F}$ der Gruppe Γ in $\mathfrak{T}$. Mit einem Akzent soll, wie in der Topologie üblich, der Rand bezeichnet werden: $\mathfrak{F}'$ ist also der Rand von $\mathfrak{F}$. Es ist

$$I(\varphi, \psi\,T) = \frac{1}{2\pi i} \sum_{j=1}^{d} \int_{\mathfrak{F}'} \Phi(\tau)\,\psi(v_j\tau) \left(\frac{a_{v_j} d_{v_j} - b_{v_j} c_{v_j}}{(c_{v_j}\tau + d_{v_j})^2} \right)^{\frac{n-2}{2}} d v_j \tau .$$

Wie in § 3 ausgeführt wurde, hängt $I(\varphi, \psi\,T)$ weder von der Integrationskonstanten des Integrals $\Phi(\tau)$ von $\varphi(\tau)$ noch von der Wahl von $\mathfrak{F}$ ab. Man setze nun $v_j = v_0 \varepsilon_j$ gemäß (34) und substituiere in den einzelnen Integralen $\varepsilon_j \tau = \tau_j$, schreibe aber wieder τ an Stelle von τ_j. Das ergibt wegen (13) und

$$\frac{a_{v_j} d_{v_j} - b_{v_j} c_{v_j}}{(c_{v_j}\tau + d_{v_j})^2} = \frac{a_{v_0} d_{v_0} - b_{v_0} c_{v_0}}{(c_{v_0}\tau + d_{v_0})^2}\,[\varepsilon_j]^{-2}, \qquad \Omega_{\varepsilon_j^{-1}}[\varepsilon_j]^{n-2} = -\Omega_{\varepsilon_j}:$$

$$I(\varphi, \psi\,T) = \frac{1}{2\pi i} \int_{\Sigma \varepsilon_j \mathfrak{F}'} \Phi(\tau)\,\psi(\tau)\,[v_0]^{-n}\,d v_0 \tau - \frac{1}{2\pi i} \sum_{j=1}^{d} \int_{\varepsilon_j \mathfrak{F}'} \Omega_{\varepsilon_j}(\tau)\,\psi(\tau)\,[v_0]^{-n}\,d\tau .$$

Die letzteren Integrale kann man durch $(n-1)$-fache Teilintegration umformen. Sie fallen sämtlich gleich 0 aus, da die $\Omega_{\varepsilon_j}(\tau)$ Polynome $(n-2)$-ten Grades sind und das $(n-1)$-fache unbestimmte Integral von $\psi(\tau)\,[v_0]^{-n}$ wegunabhängig ist, denn $\psi(\tau)$ wurde ja als eine automorphe Form 1. oder 2. Gattung vorausgesetzt. Also gilt

$$I(\varphi, \psi\,T) = \frac{1}{2\pi i} \int_{\Sigma \varepsilon_j \mathfrak{F}'} \Phi(\tau)\,\psi(\tau)\,[v_0]^{-n}\,d v_0 \tau .$$

Eine erneute Variablensubstitution $\nu_0 \tau = \sigma$ liefert

$$I(\varphi, \psi T) = \frac{1}{2\pi i} \int_{\nu_0 \cdot \sum \varepsilon_j \mathfrak{F}} \Phi(\tau) [\nu_0^{-1}]^{n-2} \psi(\tau) \, d\tau. \tag{42}$$

Jetzt ist $\sum \varepsilon_i \mathfrak{F}$ ein Fundamentalbereich der Gruppe $\Gamma \cap \nu_0^{-1} \Gamma \nu_0$ und folglich $\nu_0 \cdot \sum \varepsilon_j \mathfrak{F}$ ein Fundamentalbereich der Gruppe $\nu_0 (\Gamma \cap \nu_0^{-1} \Gamma \nu_0) \nu_0^{-1} = \Gamma \cap \nu_0 \Gamma \nu_0^{-1}$. Ferner ist $\Phi(\tau) [\nu_0^{-1}]^{n-2}$ ein Integral 1. oder 2. Gattung für die Gruppe $\nu_0 \Gamma \nu_0^{-1}$. Daher ist (42) die Bilinearform $I(\varphi [\nu_0^{-1}]^{-n}, \psi)$ bezüglich der Gruppe $\bar{\Gamma} = \Gamma \cap \nu_0 \Gamma \nu_0^{-1}$. Ein anderer Fundamentalbereich derselben Gruppe ist aber wegen (36) $\sum_{j=1}^{d^*} \eta_j \mathfrak{F}$. Daher kann man für (42) auch schreiben:

$$I(\varphi, \psi T) = \frac{1}{2\pi i} \sum_{j=1}^{d^*} \int_{\eta_j \mathfrak{F}} \Phi(\tau) [\nu_0^{-1}]^{n-2} \psi(\tau) \, d\tau.$$

Hier substituiere man in den einzelnen Integralen $\tau = \eta_j \tau_j$, schreibe aber wieder τ an Stelle von τ_j und erhält

$$I(\varphi, \psi T) = \sum_{j=1}^{d^*} \int_{\mathfrak{F}} \Phi(\tau) [\nu_0^{-1} \eta_j]^{n-2} \psi(\tau) \, d\tau.$$

Das ist aber wegen (37) und wegen der Vertauschbarkeit von Integration und Anwendung des Operators T^* gleich $I(\varphi T^*, \psi)$, wie behauptet wurde.

§ 6. Die Spur einer Korrespondenz

1. Wir greifen jetzt die in § 5, Nr. 4 gestellte Aufgabe an, die Spur der Darstellung einer Korrespondenz T in bezug auf den Modul der automorphen Formen 1. Gattung des Grades $-n$ zu bestimmen. Es sei $\varphi_i(\tau)$ $(i = 1, \ldots, G_n)$ eine Basis dieses Moduls und

$$\varphi_i(\tau) \, T = \sum_{j=1}^{G_n} m_{ij} \varphi_j(\tau), \tag{43}$$

dann ist also

$$s_n(T) = \sum_{i=1}^{G_n} m_{ii} \tag{44}$$

zu berechnen.

Seien $\psi_i(\tau)$ die zu den $\varphi_i(\tau)$ gehörigen automorphen Formen 2. Gattung der Art, daß die Gln. (29) bestehen, und $\Psi_i(\tau)$ deren Integrale. Mit diesen bilde man die Funktion

$$K_1(\tau, \sigma) = \sum_{i=1}^{G_n} \varphi_i(\tau) \Psi_i(\sigma). \tag{45}$$

Gemäß § 4, Nr. 1 wird hierzu eine automorphe Form $K_0(\sigma) = K_0(\tau, \sigma)$ des Grades $n-2$ in σ bestimmt, welche Pole an den singulären Stellen der $\Psi_i(\tau)$ mit denselben Hauptteilen wie $K_1(\tau, \sigma)$ besitzt und außer an diesen Stellen höchstens noch an der Stelle $\sigma = \tau$ (und den hierzu bezüglich Γ äquivalenten)

singulär wird; und zwar soll dort höchstens ein Pol der Ordnung G_n liegen. Mit diesen Funktionen wird

$$K(\tau,\sigma) = -K_1(\tau,\sigma) + K_0(\tau,\sigma) \tag{46}$$

gebildet. $K(\tau,\sigma)$ hat die folgenden Eigenschaften:

1) *Für jede automorphe Form $\varphi(\tau)$ von 1. Gattung gilt*

$$\frac{1}{2\pi i}\int_{\mathfrak{R}} K(\tau,\sigma)\,\varphi(\sigma)\,d\sigma = \varphi(\tau). \tag{47}$$

2) *$K(\tau,\sigma)$ ist ein Integral 2. Gattung in σ, welches überall bis auf die Stelle $\sigma=\tau$ (und die hierzu bezüglich Γ äquivalenten Stellen) holomorph ist. In $\sigma=\tau$ liegt ein Pol erster Ordnung mit dem Residuum* 1:

$$K(\tau,\sigma) = \frac{1}{\sigma-\tau} + \cdots. \tag{48}$$

Die Gl. (47) folgt aus dem Ansatz (45), (46) von $K(\tau,\sigma)$, sowie aus § 3, Nr. 2. Die Holomorphie von $K(\tau,\sigma)$ bis auf die Stelle $\sigma=\tau$ ist ebenfalls eine Folge des Ansatzes. Um (48) zu beweisen, gehe man davon aus, daß $K(\tau,\sigma)$ höchstens einen Pol G_n-ter Ordnung haben kann. Es ist also

$$K(\tau,\sigma) = \sum_{\mu=1}^{G_n} \frac{c_\mu}{(\sigma-\tau)^\mu} + \cdots.$$

Aus (47) folgt jetzt

$$\frac{1}{2\pi i}\int_{\mathfrak{R}} K(\tau,\sigma)\,\varphi(\sigma)\,d\sigma = \varphi(\tau) = \sum_{\mu=1}^{G_n} \frac{c_\mu}{(\mu-1)!}\left(\frac{d}{d\tau}\right)^{\mu-1}\varphi(\tau).$$

Diese Gleichung kann aber für alle G_n Formen $\varphi(\tau)$ 1. Gattung nur dann bestehen, wenn alle $c_\mu = 0$ sind bis auf $c_1 = 1$.

2. Zur Berechnung der Spur (44) braucht man noch den folgenden Hilfssatz: *$K_0(\tau,\sigma)$ ist in τ eine automorphe Form des Grades $-n$, wenn $n>2$ ist.*

Zum Beweise geht man davon aus, daß $K(\tau,\sigma)$ konstruktionsgemäß nicht nur in $\sigma=\tau$, sondern auch in allen hierzu bezüglich Γ äquivalenten Punkten Pole besitzt. Ferner ist $K(\tau,\sigma)$ in σ ein Integral 2. Gattung. Folglich erhält man den Hauptteil der Singularität von $K(\tau,\sigma)$ in $\sigma=\varepsilon\tau$ für ein $\varepsilon\in\Gamma$ durch Anwendung des Operators $[\varepsilon^{-1}]_\sigma^{n-2}$ bezüglich der Variablen σ auf die Entwicklung (48):

$$K(\tau,\sigma)\,[\varepsilon^{-1}]_\sigma^{n-2} = \frac{1}{(a_{\varepsilon^{-1}}\sigma + b_{\varepsilon^{-1}}) - \tau\,(c_{\varepsilon^{-1}}\sigma + d_{\varepsilon^{-1}})}\,\frac{(c_{\varepsilon^{-1}}\sigma + d_{\varepsilon^{-1}})^{n-1}}{(a_{\varepsilon^{-1}}d_{\varepsilon^{-1}} - b_{\varepsilon^{-1}}c_{\varepsilon^{-1}})^{\frac{n-2}{2}}} + \cdots$$

$$= \frac{1}{\sigma-\varepsilon\tau}\,\frac{(a_\varepsilon d_\varepsilon - b_\varepsilon c_\varepsilon)^{\frac{n}{2}}}{(c_\varepsilon\tau + d_\varepsilon)^n} + \cdots,$$

die Punkte bedeuten jedesmal eine Potenzreihe in σ. Dann ist in verständlicher Abkürzung $K(\tau,\sigma)\,[\varepsilon^{-1}]_\sigma^{n-2} - K(\tau,\sigma)\,[\varepsilon]_\tau^{-n}$ als Funktion von σ an der

Stelle $\sigma = \tau$ und folglich überall holomorph. Andererseits folgt aus der Konstruktion

$$K(\tau, \sigma)\,[\varepsilon^{-1}]_\sigma^{n-2} = K(\tau, \sigma) - \sum_{i=1}^{G_n} \varphi_i(\tau)\,\Omega_i'(\sigma),$$

wo die $\Omega_i'(\sigma)$ die Perioden der $\Psi_i(\sigma)$ sind. Mithin ist wegen (46):

$$K_0(\tau, \sigma) - K_0(\tau, \sigma)\,[\varepsilon]_\tau^{-n}$$

eine holomorphe automorphe Form in σ des Grades $n-2$. Eine solche muß bekanntlich identisch 0 sein, da $n > 2$ vorausgesetzt wurde. Das war aber die Behauptung.

3. Wir halten bis zum Schluß von Nr. 6 an der Voraussetzung $n > 2$ fest. Für eine Korrespondenz T folgt aus (43), (44)

$$s_n(T) = \sum_{j=1}^{G_n} I(\varphi_j T, \psi_j) = \frac{-1}{2\pi i} \sum_{j=1}^{G_n} \int_{\mathfrak{R}} (\varphi_j(\tau)\, T)\, \Psi_j(\tau)\, d\tau, \tag{49}$$

das Integral im Sinne von § 3, Nr. 1 verstanden.

Nun war in Nr. 2 bewiesen worden, daß $K_0(\tau, \sigma)$ in τ und σ eine automorphe Form der Grade $-n$ bzw. $n-2$ ist. Dasselbe gilt dann auch für

$$K_0(\tau, \sigma)\, T_\tau = \sum_{i=1}^{d} K_0(\nu_i \tau, \sigma)\, \frac{(a_{\nu_i} d_{\nu_i} - b_{\nu_i} c_{\nu_i})^{\frac{n}{2}}}{(c_{\nu_i}\tau + d_{\nu_i})^n},$$

und daher ist

$$K_{0T}(\tau) = [K_0(\tau, \sigma)\, T_\tau]_{\sigma=\tau}$$

eine automorphe Form des Grades -2. Die Residuensumme des Differentials $K_{0T}(\tau)\,d\tau$ ist gleich 0:

$$\int_{\mathfrak{R}} K_{0T}(\tau)\, d\tau = 0. \tag{50}$$

Aus (45), (46) einerseits, (49) und (50) andererseits ergibt sich nun

$$s_n(T) = \frac{1}{2\pi i} \int_{\mathfrak{R}} K_T(\tau)\, d\tau \quad \text{mit} \quad K_T(\tau) = \sum_{i=1}^{d} K(\nu_i \tau, \tau)\, \frac{(a_{\nu_i} d_{\nu_i} - b_{\nu_i} c_{\nu_i})^{\frac{n}{2}}}{(c_{\nu_i}\tau + d_{\nu_i})^n}. \tag{51}$$

Das Integral ist wie bisher im Sinne von § 3 als die Summe der Residuen des Integranden, einschließlich der Ecken und Spitzen zu verstehen.

Es ist wichtig zu bemerken, daß (51) auch dann gilt, wenn $G_n = 0$ ist, wenn also keine automorphen Formen 1. oder 2. Gattung existieren. Dann ist einerseits $s_n(T) = 0$. Andererseits gibt es eine automorphe Form $K(\tau, \sigma) = K_0(\tau, \sigma)$ in σ vom Grade $n-2$, welche in $\sigma = \tau$, und auch nur hier, einen Pol mit dem Hauptteil (48) hat. $K(\tau, \sigma)$ ist in τ eine automorphe Form des Grades $-n$, und die wie oben gebildete Funktion $K_T(\tau)$ ist eine automorphe Form des Grades -2, ihr Integral (51) ist daher 0.

4. Es bleibt die Auswertung von (51) übrig. $K_T(\tau)$ hat an den Stellen $\tau = \tau_0$ Pole 1. Ordnung, für welche $\nu_i \tau_0 = \eta^{-1} \tau_0$ für irgendeinen Index i aus

der Reihe $1, \ldots, d$ und irgendein $\eta \in \Gamma$ ist. Das heißt, τ_0 ist ein *Fixpunkt* von $\eta \nu_i$. Mit irgendeinem $\varepsilon \in \Gamma$ ist $\varepsilon^{-1}\tau_0$ Fixpunkt von $\varepsilon^{-1}\eta\nu_i\varepsilon$, d.h. also: zu konjugierten Substitutionen gehören äquivalente Fixpunkte. Alle in dieser Weise auftretenden Fixpunkte heißen die *Fixpunkte der Korrespondenz*. Faßt man sie in Klassen $\Gamma\tau_0$ zusammen und deutet diese als die Punkte von $\mathfrak{R}$ (§ 1, Nr. 1), so gibt es natürlich nur endlich viele Fixpunkte von T in $\mathfrak{R}$.

Sei nun $\nu = \eta\nu_i$ und $\nu\tau_0 = \tau_0$ ein Fixpunkt von T. Zunächst werde angenommen, daß τ_0 nicht eine Ecke oder Spitze ist. Der Pol von $K_T(\tau)$ in $\tau = \tau_0$ hat dann den Hauptteil

$$(52) \qquad \frac{1}{\tau - \nu\tau} \frac{(a_\nu d_\nu - b_\nu c_\nu)^n}{(c_\nu \tau + d_\nu)^n} + \cdots = \frac{1}{\tau - \tau_0} \frac{1}{1 - \dfrac{a_\nu d_\nu - b_\nu c_\nu}{(c_\nu \tau_0 + d_\nu)^2}} \frac{(a_\nu d_\nu - b_\nu c_\nu)^n}{(c_\nu \tau_0 + d_\nu)^n}.$$

Das Residuum kann folgendermaßen gedeutet werden: Man bezeichne die beiden Eigenwerte der Matrix ν mit

$$(53) \qquad \begin{cases} \varrho_\nu = \dfrac{a_\nu + d_\nu + \sqrt{(a_\nu + d_\nu)^2 - 4(a_\nu d_\nu - b_\nu c_\nu)}}{2} = c_\nu \tau_0 + d_\nu, \\ \varrho_\nu^* = \dfrac{a_\nu + d_\nu - \sqrt{(a_\nu + d_\nu)^2 - 4(a_\nu d_\nu - b_\nu c_\nu)}}{2} = \dfrac{a_\nu d_\nu - b_\nu c_\nu}{\varrho_\nu}. \end{cases}$$

In dieser Bezeichnung ist das Residuum von $K_T(\tau)$ an dieser Stelle

$$(54) \qquad \operatorname{Res}(\nu) = \frac{(a_\nu d_\nu - b_\nu c_\nu)^n}{\varrho_\nu^{n-1}} \frac{1}{\varrho_\nu - \varrho_\nu^*} = \frac{1}{(a_\nu d_\nu - b_\nu c_\nu)^{\frac{n-2}{2}}} \frac{\varrho_\nu^{*n-1}}{\varrho_\nu - \varrho_\nu^*}.$$

Zum richtigen Verständnis von (54) ist eine Unterscheidung der beiden Eigenwerte (53) von ν erforderlich. Zunächst sieht man aus (54), daß $\varrho_\nu \neq \varrho_\nu^*$ ist. Das Vorzeichen der Quadratwurzel in (53) ist so zu nehmen, daß $\tau_0 = \dfrac{\varrho_\nu - d_\nu}{c_\nu}$ im Inneren des Bereichs $\mathfrak{T}$ liegt, auf den die Funktion τ die Riemannsche Fläche $\mathfrak{R}$ abbildet. Wenn τ_0 für beide Vorzeichen der Quadratwurzel in $\mathfrak{T}$ liegt, gehören zu diesem ν zwei Pole an den Stellen τ_0, τ_0^*, es sei denn, daß $\tau_0^* = \varepsilon\tau_0$ mit einem $\varepsilon \in \Gamma$ gilt.

Ist τ_0 eine Ecke $\mathfrak{e}_\lambda$, so behält das Residuum von $R_T(\tau)$ als Funktion in τ den Wert (54). Da aber die lokale Uniformisierende $z = (\tau - \tau_0)^{n_\lambda}$ ist, gilt $z^{-1}dz = n_\lambda(\tau - \tau_0)^{-1}d(\tau - \tau_0)$, und es gilt hier bis auf ein in τ_0 holomorphes Differential

$$K_T(\tau)\,d\tau = \operatorname{Res}(\nu) \frac{d\tau}{\tau - \tau_0} = \frac{\operatorname{Res}(\nu)}{n_\lambda} \frac{dz}{z}.$$

Man muß daher (54) durch n_λ dividieren, um den entsprechenden Beitrag zu (51) zu erhalten.

Endlich sei der Fixpunkt von ν eine Spitze $\mathfrak{s}$. Wir denken uns diesen Punkt nach $\tau_0 = i\infty$ verlegt, er ist dann gleichzeitig der Fixpunkt des Elements $\gamma = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \in \Gamma$, welches wir gemäß § 1 in eine bequeme Normalform

transformiert annehmen. Es muß dann

$$\nu = \begin{pmatrix} a_\nu & b_\nu \\ 0 & d_\nu \end{pmatrix} \tag{55}$$

gelten, wobei $a_\nu d_\nu$ reell und >0 ist. Wegen

$$\nu^{-1}\gamma\nu = \begin{pmatrix} 1 & a_\nu^{-1} d_\nu \\ 0 & 1 \end{pmatrix}$$

können die Gruppenindizes $[\Gamma : \Gamma \cap \nu^{-1}\Gamma\nu]$ und $[\nu^{-1}\Gamma\nu : \Gamma \cap \nu^{-1}\Gamma\nu]$ nur dann endlich sein, wenn $a_\nu^{-1} d_\nu$ rational ist. Wir dürfen nun ohne Beschränkung der Allgemeinheit voraussetzen, daß a_ν und d_ν teilerfremde natürliche Zahlen sind.

Wir brauchen ferner einige der Elemente (34), nämlich

$$\nu_i = \nu\gamma^{i-1} = \begin{pmatrix} a_\nu & b_\nu + a_\nu(i-1) \\ 0 & d_\nu \end{pmatrix}, \qquad i = 1, \ldots, d_\nu. \tag{56}$$

Für $i > d_\nu$ ist $\nu\gamma^{i-1}$ in der Form $\gamma^j \nu_k$ für gewisse j und k ausdrückbar. Die übrigen ν_i in (34), für welche ε_i nicht eine Potenz von γ ist, haben nicht mehr denselben Fixpunkt $\mathfrak{s}$ und interessieren deshalb in diesem Zusammenhang nicht.

In der lokalen Uniformisierenden $z = e^{2\pi i \tau}$ schreibt sich $K(\tau, \sigma)$ nach (48) so:

$$K(\tau, \sigma) = 2\pi i \frac{e^{2\pi i \tau}}{e^{2\pi i \sigma} - e^{2\pi i \tau}} + \cdots.$$

Und es ist

$$K(\tau, \sigma)\, T_\tau = \sum_{i=1}^{d_\nu} K(\nu_i \tau, \sigma) \frac{\left(a_{\nu_i} d_{\nu_i} - b_{\nu_i} c_{\nu_i}\right)^{\frac{n}{2}}}{\left(c_{\nu_i}\tau + d_{\nu_i}\right)^n} + \cdots,$$

wo die Punkte eine Potenzreihe in $e^{2\pi i \sigma}$ andeuten. Die Ausrechnung ergibt

$$\begin{aligned} K(\tau, \sigma)\, T_\tau &= -\frac{d \log\left(e^{2\pi i d_\nu \sigma} - e^{2\pi i (a_\nu \tau + b_\nu)}\right)}{d\, \frac{a_\nu \tau + b_\nu}{d_\nu}} \left(\frac{a_\nu}{d_\nu}\right)^{\frac{n}{2}} + \cdots \\ &= 2\pi i \frac{a_\nu^{n-1}}{(a_\nu d_\nu)^{\frac{n-2}{2}}} \frac{e^{2\pi i (a_\nu \tau + b_\nu)}}{e^{2\pi i d_\nu \sigma} - e^{2\pi i (a_\nu \tau + b_\nu)}} + \cdots. \end{aligned}$$

Das Residuum von $K_T(\tau)\, d\tau$ in $\tau = i\infty$ ist nun

$$\text{Res}(\nu) = \begin{cases} 0, & a_\nu > d_\nu, \\ -\dfrac{a_\nu^{n-1}}{\left(1 - \zeta(\nu)\, e^{-2\pi i b_\nu}\right)(a_\nu d_\nu)^{\frac{n-2}{2}}}, & a_\nu \leq d_\nu, \end{cases} \qquad \zeta(\nu) = \begin{cases} 1, & a_\nu = d_\nu, \\ 0, & a_\nu < d_\nu. \end{cases} \tag{57}$$

Jetzt läßt sich die Berechnung der Spur (51) in dem folgenden Rezept zusammenfassen:

Die Spur einer Korrespondenz (1. Formulierung). *Man bilde alle Matrizen $\nu = \varepsilon_1 \nu_0 \varepsilon_2$, ε_1 und ε_2 beliebig aus Γ, welche Fixpunkte τ_0 im Inneren oder auf dem Rande von $\mathfrak{T}$ haben, und zwar nehme man aus jeder Klasse*

konjugierter Elemente $\varepsilon^{-1}\nu\varepsilon$ je einen Vertreter. Die Spur von T als Operator im Modul aller automorphen Formen 1. Gattung eines geraden Grades $-n<-2$ ist dann die über ν zu erstreckende Summe

$$s_n(T)=\sum_{\nu}\frac{1}{n_\nu}\operatorname{Res}(\nu), \tag{58}$$

deren Summanden in (54) und (57) angegeben und folgendermaßen erklärt werden.

Hat ν zwei verschiedene Eigenwerte ϱ_ν, ϱ_ν^*, so denke man sich ν als Matrix $\nu=\begin{pmatrix}a_\nu & b_\nu\\ c_\nu & d_\nu\end{pmatrix}$ mit $c_\nu\neq 0$, was die Allgemeinheit nicht beeinträchtigt Ein Fixpunkt von diesem ν ist dann $\tau_0=\frac{\varrho_\nu-d_\nu}{c_\nu}$. In (54) hat man denjenigen Eigenwert zu nehmen, für den der so berechnete Fixpunkt τ_0 im Inneren von $\mathfrak{T}$ liegt. Liegt dagegen τ_0 für beide Eigenwerte ϱ_ν, ϱ_ν^* in $\mathfrak{T}$, so hat man entsprechend zwei Summanden in (58) einzusetzen, es sei denn, daß die beiden Fixpunkte τ_0, $\tau_0^*=\frac{\varrho_\nu^*-d_\nu}{c_\nu}$ bezüglich Γ äquivalent sind.

Die Zahl n_ν ist im allgemeinen gleich 1 zu nehmen, wenn der Fixpunkt keine Ecke $\mathfrak{e}_\lambda$ ist, sonst gleich der Ordnungszahl dieser Ecke.

Hat ν nur einen einzigen Eigenwert, so kann man ν in die Normalform (55) transformieren, wobei a_ν und d_ν teilerfremde natürliche Zahlen sind. Zu einem solchen ν gehört der Beitrag (57).

Endlich stimmen die Spuren adjungierter Korrespondenzen überein:

$$s_n(T)=s_n(T^*).$$

Letzteres geht aus deren Definition (35) und (37) hervor.

5. Wir wollen das Berechnungsverfahren für die Spur von T noch anders formulieren, dabei aber die vereinfachende Annahme machen, daß keine Ecken und Spitzen vorhanden sind. Wegen (52) kann man (54) auch so schreiben:

$$\operatorname{Res}(\nu)=\left[\frac{1}{1-\frac{d\nu\tau}{d\tau}}\left(\frac{d\nu\tau}{d\tau}\right)^{n/2}\right]_{\tau=\tau_0}.$$

Dieser Ausdruck läßt sich noch anders schreiben. Es sei z eine in einem Punkte $\mathfrak{p}$ von $\mathfrak{R}$ bzw. $\tau=\tau_0$ verschwindende lokale Uniformisierende. Ein Zweig der Korrespondenz T bildet nun die Umgebung von $z=0$ auf sich umkehrbar eindeutig und konform ab. Die Abbildungsfunktion $T(z)$ habe hier die Reihenentwicklung $T(z)=v_\mathfrak{p}z+v_\mathfrak{p}'z^2+\cdots$. Der Koeffizient $v_\mathfrak{p}$ ist nun ersichtlich invariant bei Übergang zu einer anderen Uniformisierenden; $v_\mathfrak{p}$ soll das *Verzerrungsverhältnis* heißen. Nimmt man speziell $z=\tau-\tau_0$, so ist

$$v_\mathfrak{p}=\left[\frac{dT(z)}{dz}\right]_{z=0}=\left[\frac{d\nu\tau}{d\tau}\right]_{\tau=\tau_0}.$$

Die Spur einer Korrespondenz (2. Formulierung). *Wenn keine Ecken und Spitzen vorhanden sind, so ist die Spur einer Korrespondenz T als Operator im Modul der automorphen Formen 1. Gattung eines geraden Grades $-n<-2$*

die über alle Fixpunkte $\mathfrak{p}$ *in* $\mathfrak{R}$ *erstreckte Summe*

$$s_n(T) = \sum_{\mathfrak{p}} \frac{1}{1 - v_{\mathfrak{p}}} v_{\mathfrak{p}}^{n/2}, \tag{59}$$

wo $v_{\mathfrak{p}}$ *das soeben erklärte Verzerrungsverhältnis im Punkte* $\mathfrak{p}$ *bedeutet.*

Dieses Verzerrungsverhältnis ist bei den hier betrachteten Korrespondenzen also niemals gleich 1. Man muß allerdings im Auge behalten, daß wir hier nicht die allgemeinsten Korrespondenzen studieren.

6. Die erhaltene Spurformel (58) gestattet noch eine Ergänzung und Umformung. Man wende die Gl. (41) in (49) an und kommt dann auf

$$s_n(T) = \frac{1}{2\pi i} \int_{\mathfrak{R}} K_T^*(\tau)\, d\tau, \qquad K_T^*(\tau) = \sum_{j=1}^{d^*} K(\tau, \nu_j^* \tau) \frac{(c_{\nu_j^*} \tau + d_{\nu_j^*})^n}{(a_{\nu_j^*} d_{\nu_j^*} - b_{\nu_j^*} c_{\nu_j^*})^{n/2}},$$

wo die ν_j^* die in § 5, Nr. 2 erklärte Bedeutung haben. Bei der Berechnung der Pole von $K_T^*(\tau)$ kommen an Stelle der Substitutionen $\nu = \varepsilon_1 \nu_0 \varepsilon_2$ Substitutionen $\nu^* = \varepsilon_1 \nu_0^{-1} \varepsilon_2$ vor. Da ein gemeinsamer Faktor aller Koeffizienten ohne Einfluß ist, kann man für die ν^* gerade die Substitutionen $\nu^* = \begin{pmatrix} s(\nu) & 0 \\ 0 & s(\nu) \end{pmatrix} - \nu$ nehmen ($s(\nu) =$ Spur von ν). Diese haben die gleichen Eigenwerte wie ν. Das Residuum von $K_T^*(\tau)$ an einer Stelle τ_0, welche Fixpunkt eines solchen ν^* ist, ist an Stelle von (54) jetzt

$$\operatorname{Res}(\nu^*) = \frac{1}{(a_\nu d_\nu - b_\nu c_\nu)^{\frac{n-2}{2}}} \frac{\varrho_\nu^{n-1}}{\varrho_\nu^* - \varrho_\nu},$$

wenn $\varrho_\nu \neq \varrho_\nu^*$ war. Fallen dagegen die Eigenwerte von ν zusammen, so überlegt man sich im Anschluß an Nr. 3 mühelos, daß das Residuum von $K_T^*(\tau)$ in $\tau = i\infty$ dasselbe wie das von $K_T(\tau)$ ist.

Die Spur läßt sich demnach auch in der Gestalt

$$s_n(T) = -\frac{1}{2} \sum \frac{1}{n_\nu} \frac{1}{(a_\nu d_\nu - b_\nu c_\nu)^{\frac{n-2}{2}}} \frac{\varrho_\nu^{n-1} - \varrho_\nu^{*\,n-1}}{\varrho_\nu - \varrho_\nu^*} - \sum_{a_\nu \leq d_\nu} \frac{a_\nu^{n-1}}{(1 - \zeta(\nu) e^{-2\pi i b_\nu})\,(a_\nu d_\nu)^{\frac{n-2}{2}}}$$

schreiben. Nun ist aber

$$\frac{1}{(a_\nu d_\nu - b_\nu c_\nu)^{\frac{n-2}{2}}} \frac{\varrho_\nu^{n-1} - \varrho_\nu^{*\,n-1}}{\varrho_\nu - \varrho_\nu^*} = \frac{1}{|\nu|^{\frac{n-2}{2}}} (\varrho_\nu^{n-2} + \varrho_\nu^{n-3} \varrho_\nu^* + \cdots)$$

die Spur $s([\nu]^{n-2})$ der in § 1, Nr. 2 erwähnten Darstellung $[\nu]^{n-2}$ der projektiven linearen Gruppe. Also ist endlich

$$s_n(T) = -\frac{1}{2} \sum \frac{1}{n_\nu} s([\nu]^{n-2}) - \sum_{a_\nu \leq d_\nu} \frac{a_\nu^{n-1}}{(1 - \zeta(\nu) e^{-2\pi i b_\nu})\,(a_\nu d_\nu)^{\frac{n-2}{2}}}; \tag{60}$$

wegen der Bedeutung der benutzten Symbole muß auf den Zusammenhang verwiesen werden.

7. Durch eine leichte Überlegung läßt sich die Spurformel (58), (59), (60) so korrigieren, daß sie auch noch für $n = 2$ gilt. Für $n = 2$ ist die Behauptung von Nr. 2 falsch, doch sieht man, daß $K_0(\tau, \sigma) - K_0(\tau, \sigma)\,[\varepsilon]_\tau^{-n}$ eine in σ

überall holomorphe und bei der Gruppe Γ invariante Funktion ist. Sie hängt daher nur noch von τ ab. Aus dem Grunde sind $K_0(\tau,\sigma)-K_0(\tau,\sigma_0)$ und $[K_0(\tau,\sigma)\,T_\tau]_{\sigma=\tau}-K_0(\tau,\sigma_0)\,T$ für eine beliebige Konstante σ_0 automorphe Formen in τ des Grades -2; die letztere erfüllt wieder die Gl. (50). Wie oben ergibt sich jetzt an Stelle von (51):

$$s_n(T)=\frac{1}{2\pi i}\int_{\mathfrak{R}} K_T(\tau)\,d\tau-\frac{1}{2\pi i}\int_{\mathfrak{R}} K_0(\tau,\sigma_0)\,T\,d\tau. \tag{61}$$

Wegen (48) ist

$$K_0(\tau,\sigma_0)\,T=\sum_{i=1}^{d} K_0(\nu_i\tau,\sigma_0)\,\frac{a_{\nu_i}d_{\nu_i}-b_{\nu_i}c_{\nu_i}}{(c_{\nu_i}\tau+d_{\nu_i})^2}$$

$$=\sum_{i=1}^{d}\frac{1}{\nu_i^{-1}\sigma_0-\tau}\,\frac{a_{\nu_i}d_{\nu_i}-b_{\nu_i}c_{\nu_i}}{(c_{\nu_i}\tau+d_{\nu_i})\,(-c_{\nu_i}\sigma_0+a_{\nu_i})}+\cdots=\sum_{i=1}^{d}\frac{1}{\nu_i^{-1}\sigma_0-\tau}+\cdots,$$

wo die Punkte eine Potenzreihe in τ bedeuten. Folglich hat das zweite Integral den Wert d.

Im Falle $n=2$ hat man rechts in (58) *und* (59) *noch die durch* (32) *erklärte Zahl d, in* (60) $\frac{1}{2}(d+d^*)$ *hinzuzufügen.*

8. Zum Schluß bringen wir ein Beispiel. Es sei Γ die Modulgruppe und N eine quadratfreie natürliche Zahl. Die Modularkorrespondenz $T=T_N$ läßt sich dann in der Weise von § 5 erzeugen. Der Bereich $\mathfrak{T}$ ist die obere τ-Halbebene. Die Substitutionen, über die in (60) summiert wird, sind ganzzahlige Matrizen der Determinante N, welche entweder komplexe Eigenwerte haben (erste Summe), oder welche in die Gestalt (55) mit ganzen rationalen a_ν, b_ν, d_ν unimodular transformierbar sind (zweite Summe). Die ersteren ν haben also eine Spur $s=s(\nu)$, die zwischen $-2\sqrt{N}$ und $+2\sqrt{N}$ liegt. Die Anzahl der nicht konjugierten ν mit vorgegebener Spur ist gleich der Anzahl $h(s^2-4N)$ der Ringidealklassen für den quadratischen Zahlring mit der Diskriminante s^2-4N. Die Zahl n_ν ist die halbe Einheitenanzahl $e(s^2-4N)$ dieses Ringes, also $e(-3)=3$, $e(-4)=2$, $e(s^2-4N)=1$ in allen anderen Fällen. Die Spurformel (60) wird also

$$\begin{cases} s_n(T_N)=N^{\frac{2-n}{2}}\times \\ \times\left[-\frac{1}{2}\sum_{s^2<4N}\frac{h(s^2-4N)}{e(s^2-4N)}\,\frac{\varrho_s^{n-1}-\varrho_s^{*\,n-1}}{\varrho_s-\varrho_s^*}-\sum_{\substack{t/N\\ t<\sqrt{N}}}t^{n-1}+\begin{cases}\sum t & \text{für } n=2,\\ 0 & \text{für } n>2\end{cases}\right] \end{cases} \tag{62}$$

mit $\varrho_s=\frac{1}{2}\left(s+\sqrt{s^2-4N}\right)$, $\varrho_s^*=\frac{1}{2}\left(s-\sqrt{s^2-4N}\right)$.

Es läßt sich übrigens leicht zeigen, daß (62) auch für nicht quadratfreies N gilt. Für eine Quadratzahl $N=M^2$ ist dagegen rechts in der Klammer noch $\frac{1}{2}M^{n-1}$ zu subtrahieren. Analoge Formeln für gewisse Kongruenzuntergruppen der Modulgruppe haben eine ähnliche Bauart. Die Herleitung dieser Formeln macht keine besondere Mühe, da die Anzahlen der nicht konjugierten ν mit gleicher Spur bereits in einer früheren Arbeit [2] bestimmt wurden.

§ 7. Schlußbemerkungen

1. Zum Schluß sollen einige offene Fragen besprochen werden, die durch das Vorhergehende nahegelegt werden.

Wie ordnet sich im Falle $n = 2$ das Abelsche Theorem in diesen Zusammenhang ein? Es sei $\chi(\tau)\,d\tau$ ein Differential 3. Gattung und $\varphi(\tau)\,d\tau$ ein Differential 1. oder 2. Gattung, welches in den Singularitäten des ersteren holomorph ist. Man stellt leicht mittels des Cauchyschen Integralsatzes fest

$$I(\varphi, \chi) = \sum_i \alpha_i \int_{\mathfrak{p}_i}^{\mathfrak{q}_i} \varphi(\tau)\,d\tau,$$

dabei brauchen die Punkte $\mathfrak{p}_i$, $\mathfrak{q}_i$ nicht notwendig verschieden zu sein. Die Bedeutung der Konstanten α_i ist die Folgende: Für einen Punkt $\mathfrak{p}$ bilde man die Summe $\sum \alpha_i - \sum \alpha_j = R_{\mathfrak{p}}$, erstreckt über alle $\mathfrak{p}_i$ und alle $\mathfrak{q}_j$, die mit $\mathfrak{p}$ identisch sind. Die Zahl $R_{\mathfrak{p}}$ ist dann das Residuum des Differentials $\chi(\tau)\,d\tau$ an der Stelle $\mathfrak{p}$. Die die Punkte $\mathfrak{p}_i$ und $\mathfrak{q}_i$ verbindenden Wege sind in der aufgeschnittenen Fläche $\mathfrak{R}$ zu wählen. Das Integral $I(\varphi, \chi)$ ist von der Art der Aufschneidung zu einem (oder mehreren) einfach zusammenhängenden Flächenstück abhängig. Setzt man voraus, daß $\chi(\tau)\,d\tau$ nur Pole mit ganzzahligen Residuen hat, so sind die α_i auch ganze Zahlen, und $I(\varphi, \chi)$ ist als Restklasse modulo den Perioden von $\int \varphi(\tau)\,d\tau$ eindeutig definiert. Es gilt folgende

Modifikation des Abelschen Theorems. *Ein Differential $\chi(\tau)\,d\tau$ mit ganzen rationalen Residuen ist dann und nur dann das Differential des Logarithmus einer meromorphen Funktion auf $\mathfrak{R}$, wenn für alle Differentiale $\varphi(\tau)\,d\tau$ 1. und 2. Gattung $I(\varphi, \chi) \equiv 0$ modulo den Perioden von $\int \varphi(\tau)\,d\tau$ ist.*

2. In der komplexen Multiplikation der singulären elliptischen Funktionen spiegeln sich wichtige Eigenschaften der Riemannschen Fläche $\mathfrak{R}$ in dem Modul der Perioden der gewöhnlichen Differentiale 1. Gattung. Gibt es Analoga dieses Sachverhalts, bei welchen die Perioden der Differentiale höheren Grades in Erscheinung treten? Diese Frage wird dadurch nahegelegt, daß nach § 2, Nr. 4 die Perioden bei geeigneter Wahl der Differentiale in dem Körper liegen, dem die Koeffizienten der Gruppe Γ angehören. Dieses ist bei Differentialen höheren Grades, deren Perioden ja Polynome sind, nicht trivial. Auch der Umstand, daß die Koeffizienten der bilinearen Periodenrelationen bis auf den gemeinsamen Faktor $(2\pi i)^{-1}$ ganz rational sind, legt es nahe, nach zahlentheoretischen Anwendungen zu suchen.

3. Wegen der eineindeutigen Zuordnung (§ 4, Nr. 3) von automorphen Formen 1. und 2. Gattung auf der einen Seite und Cohomologieklassen auf der anderen, ferner wegen der Möglichkeit (40), die Wirkung der Korrespondenzen auf die Cohomologieklassen direkt zu studieren, müßte die Spurformel (60) algebraisch beweisbar sein. Bekanntlich gestattet die für $n = 2$ gültige Modifikation von (60) (s. § 6, Nr. 7) in Gestalt des Lefschetzschen Fixpunktsatzes einen rein topologischen Beweis (s. hierzu [4]). Für $n > 2$ wäre die Homologietheorie durch die Cohomologietheorie zu ersetzen.

4. Es wurde bereits gesagt, daß sich unsere Überlegungen einem allgemeineren Gedanken von A. WEIL unterordnen. Das geschieht in folgender Weise. Es sei $\mathfrak{t}(\tau)$ eine Basis des Moduls der Polynome $(n-2)$-ten Grades in τ, als ein Spaltenvektor geschrieben. Es gilt dann mit einer Matrix $M(\varepsilon)$

$$M(\varepsilon)\,\mathfrak{t}(\tau) = \mathfrak{t}(\tau)\,[\varepsilon]^{n-2}.$$

Ist $\varphi(\tau)$ eine automorphe Form $(-n)$-ten Grades und 1. oder 2. Gattung, so genügt der Formenvektor $\mathfrak{f}(\tau) = \mathfrak{t}(\tau)\,\varphi(\tau)$ den Funktionalgleichungen

$$M(\varepsilon)^{-1}\mathfrak{f}(\tau)\,[\varepsilon]^{-2} = \mathfrak{f}(\tau), \qquad \varepsilon \in \Gamma. \tag{62}$$

Anstatt dessen kann man auch sagen: Das Differential $\mathfrak{f}(\tau)\,d\tau$ ist bei der Substitution ε und gleichzeitiger Multiplikation mit $M(\varepsilon)^{-1}$ invariant:

$$M(\varepsilon)^{-1}\mathfrak{f}(\varepsilon\tau)\,d\varepsilon\tau = \mathfrak{f}(\tau)\,d\tau, \qquad \varepsilon \in \Gamma.$$

Das Integral

$$\mathfrak{F}(\tau) = \int_{\tau_0}^{\tau} \mathfrak{f}(\tau)\,d\tau \tag{63}$$

ist wegunabhängig. Bei Variablentransformation verhält es sich so:

$$M(\varepsilon)^{-1}\mathfrak{F}(\varepsilon\tau) = \mathfrak{F}(\tau) + \mathfrak{d}_\varepsilon, \qquad \varepsilon \in \Gamma,$$

wobei der (konstante) Periodenvektor $\mathfrak{d}_\varepsilon$ offenbar in einfacher Weise aus den Koeffizienten des Polynoms $\Omega_\varepsilon(\tau)$ berechnet wird. Die $\mathfrak{d}_\varepsilon$ genügen der zu (16) analogen Gleichung

$$\mathfrak{d}_{\varepsilon\eta} = M(\eta)^{-1}\,\mathfrak{d}_\varepsilon + \mathfrak{d}_\eta. \tag{64}$$

Die Theorie der Integrale (12) ist völlig gleichbedeutend mit der der Ideale (63). Letztere sind aber allgemeiner, da man sie für eine willkürliche Darstellung $M(\varepsilon)$ von Γ bilden kann. Man geht dabei von solchen Formenvektoren $\mathfrak{f}(\tau)$ aus, welche (62) mit dieser Darstellung erfüllen. Die Anzahlen der linear unabhängigen Differentiale $\mathfrak{f}(\tau)\,d\tau$ bekommt man aus WEILS Verallgemeinerung des Riemann-Rochschen Satzes. Mit diesem hat man ein Werkzeug in der Hand, die Theorie der Integrale 1. und 2. Gattung im allgemeinsten Falle soweit zu führen, wie es hier in einem speziellen geschah.

Literatur

[1] BOL, G.: Invarianten linearer Differentialgleichungen. Abh. math. Seminar Hamburger Univ. **16**, 1—28 (1949). — [2] EICHLER, M.: Über die Darstellbarkeit von Modulformen durch Thetareihen. J. reine u. angew. Math. **195**, 156—171 (1956). — [3] EICHLER, M.: On the class number of imaginary quadratic fields and the sums of divisors of natural numbers. J. Indian Math. Soc. **19**, 153—180 (1955). — [4] EICHLER, M.: Modular correspondences and their representations. Report of an international colloquium on zeta-functions, Bombay 1956. J. Indian Math. Soc. **20**, 163—206 (1956). — [5] PETERSSON, H.: Konstruktion der Modulformen ... von positiver reeller Dimension ... Sitzgsber. Akad. Wiss. Heidelberg **1950**, 417—494. — [6] WEIL, A.: Généralization des fonctions Abéliennes. J. Math. pures appl., IX. sér. **17**, 47—87 (1938). — [7] WEIL, A.: Über Matrizenringe auf Riemannschen Flächen und den Riemann-Rochschen Satz. Abh. math. Seminar Hamburger Univ. **11**, 110—115 (1936). — [8] SELBERG, A.: Harmonic analysis and discontinuous groups in weakly symmetric Riemannian spaces. Loc. cit. [4], S. 47—87.

Marburg, Mathematisches Institut der Universität

(Eingegangen am 16. Oktober 1956)

SCHMIDT, J.
Math. Zeitschr. Bd. 67, S. 299—302 (1957)

Einfacher, ordinalzahlfreier Beweis für die Wohlordnung der Mächtigkeiten

Von

JÜRGEN SCHMIDT

Unter den verschiedenen äquivalenten Formulierungen des Auswahlaxioms ist seiner Schmiegsamkeit in der Anwendung wegen das als Lemma von ZORN bekannte, letztlich auf HAUSDORFF zurückgehende Maximalprinzip[1]) in die Schlüsselposition gerückt; in der Tat lassen sich ja viele dieser Formulierungen, insbesondere das Auswahlaxiom selbst und der Zermelosche Wohlordnungssatz, als direkte Spezialisierungen des allgemeinen Maximalprinzips auffassen. Lange Zeit war hier eine gewisse Einschränkung zu machen gewesen, insofern die zahlreichen kardinalzahlarithmetischen Äquivalenzen für das Auswahlaxiom[2]), nach dem immer wieder reproduzierten Vorgang von CANTOR, weniger auf dem allgemeinen Maximalprinzip als speziell auf dem Wohlordnungssatz und der Ordinalzahltheorie zu beruhen schienen; dahingestellt blieb, ob nicht auch etwa der Satz von der Vergleichbarkeit oder gar von der Wohlordnung der Mächtigkeiten den Charakter eines speziellen Maximalprinzips besäße[3]). So hatte man sich lange Zeit damit abgefunden[4]), daß diese und andere Fundamentalsätze aus der Theorie der abstrakten (strukturlosen) Mengen anscheinend nur auf dem Umwege über eine hinreichend ausgebaute Theorie gewisser strukturierter, eben der wohlgeordneten Mengen zu gewinnen seien; lange schien es bei der etwas paradoxen Situation bleiben zu sollen, daß das allgemeine Maximalprinzip — ein für die heutige Mathematik grundlegendes und, wie es scheint, unentbehrliches logisches Prinzip — zwar seine Elastizität, seine bequeme Anwendbarkeit in reicheren Strukturtheorien wie Algebra,

[1]) Der Terminus „Maximalprinzip", obwohl vielleicht sprachlich anstößig, scheint uns sachlich angemessener als etwa „Maximumprinzip"; behauptet wird die Existenz eines maximalen Elements, nicht eines Maximums.

[2]) Siehe BACHMANN [*1*] p. 132ff., Formulierungen $(\mathfrak{A}_3)$—$(\mathfrak{A}_{12})$.

[3]) Auf eine direkte Herleitung der Vergleichbarkeit, ja der Wohlordnung der Mächtigkeiten aus dem Zornschen Lemma scheint bereits 1939 die Bemerkung von BOURBAKI [*2*] p. 39 hinzudeuten: ... en utilisant le théorème de ZORN (et par suite l'axiome de choix), on démontre ... (die zitierten Sätze); jedenfalls hat diese Bemerkung von BOURBAKI den Anstoß zu den beiden Noten [*4*] und [*5*] von FARAH gegeben. Der Beweis, den die Bourbaki-Schule 17 Jahre später ausgeführt hat, ist allerdings wieder der klassische ([*3*] p. 56, Beweis von Théorème 1).

[4]) Aber nicht ohne Unbehagen; FRAENKEL [*6*] p. 318 spricht von „the uneasiness caused by basing the comparability theorem upon the theory of order".

Topologie, Funktionalanalysis vollauf bewiesen hatte, in der allgemeinen Mengenlehre indes und damit letztlich der Logik selbst anscheinend weniger unmittelbaren Wert besaß[5]).

Den ersten Schritt zur Überwindung dieser Situation hat 1940 ZORN selbst getan; unter anderem hat er die Vergleichbarkeit der Mächtigkeiten auf die folgende, ebenso einfache wie natürliche Weise als einen Spezialfall seines berühmten Lemmas, des allgemeinen Maximalprinzips, dargestellt[6]):

Seien E und F irgendwelche Mengen; dann bezeichne $\mathsf{T}(E,F)$ die Menge aller eineindeutigen Abbildungen φ aus E, d.h. von einem Definitionsbereich $D(\varphi)\subseteq E$, in F, d.h. auf einen Wertevorrat $W(\varphi)\subseteq F$. Da diese Abbildungsmenge zumindest die leere Abbildung enthält, ist sie jedenfalls nicht leer. Sie wird (teilweise) geordnet durch die Relation $\varphi\subseteq\psi$, φ Einschränkung von ψ, d.h. $D(\varphi)\subseteq D(\psi)$ und $\varphi(x)=\psi(x)$ für jedes $x\in D(\varphi)$. Ist K eine Kette (total, d.h. einfach geordnete Teilmenge) in der so geordneten Menge $\mathsf{T}(E,F)$, so besitzt K darin eine obere Schranke, ja eine obere Grenze, nämlich die Vereinigung σ der Abbildungen $\varphi\in K$, deren Definitionsbereich $D(\sigma)$ und Wertevorrat $W(\sigma)$ übrigens gleich der Vereinigung der Definitionsbereiche $D(\varphi)$ bzw. Wertevorräte $W(\varphi)$ der $\varphi\in K$ ist. Die Ordnung von $\mathsf{T}(E,F)$ ist also induktiv[7]); nach dem allgemeinen Maximalprinzip[8]) existiert also in $\mathsf{T}(E,F)$ ein maximales Element μ. Dann ist aber $D(\mu)=E$ und damit μ eine eineindeutige Abbildung von E in F oder aber $W(\mu)=F$ und damit die Umkehrung μ^{-1} eine eineindeutige Abbildung von F in E; denn wäre weder das eine noch das andere der Fall, so ließe sich offenbar eine Transformation $\nu>\mu$ konstruieren, wider die Maximaleigenschaft von μ.

Genau dieser Beweis bildet dann 1952 den Inhalt einer Note von FARAH[9]). 1953 fügt FRAENKEL dem klassischen, Ordinalzahlen benutzenden Beweis des Vergleichbarkeitssatzes einen ordinalzahlfreien Beweis an, dessen Länge (von fast drei engbedruckten Seiten) allein dem Verzicht auf das allgemeine Maximalprinzip zuzuschreiben, dessen Kern aber mit dem Beweis von ZORN schlechthin identisch ist[10]). An den Beweis von ZORN schließlich hat 1955 SEKI ganz gewiß gedacht, als er den direkten Schluß vom allgemeinen Maximalprinzip auf den Vergleichbarkeitssatz als trivial und damit keiner Wiedergabe bedürftig bezeichnete[11]).

---

[5]) Besonders deutlich wird dieser Kontrast in den „Eléments de Mathématique" von BOURBAKI, die sich (s. Fußnote 3) gegenüber den Hauptsätzen der Kardinalzahltheorie völlig klassisch verhalten, während sie in den reicheren Theorien durch ausschließliche Verwendung des Zornschen Lemmas beträchtliche Beweisvereinfachungen erzielen und damit die Bedeutung dieses Prinzips besonders herausstellen.

[6]) ZORN [9] p. 11, (2.2).

[7]) BOURBAKI [3] p. 44ff.

[8]) BOURBAKI [3] p. 45, Théorème 2.

[9]) FARAH [4].

[10]) FRAENKEL [6] p. 319ff.

[11]) SEKI [8].

Der Zornsche Beweis für die Vergleichbarkeit läßt sich nun — und das ist der eigentliche Gegenstand dieser Mitteilung — mit geringfügigem technischem Aufwand sogar zum Beweis der Wohlordnung der Mächtigkeiten ausbauen. Daran scheint von den genannten Autoren niemand gedacht zu haben; ja FARAH leitet 1953 in einer zweiten Note die Wohlordnung der Mächtigkeiten aus dem Zornschen Lemma nur auf dem Umwege über den Wohlordnungssatz her und läßt damit praktisch alles beim alten[12]). Der direkte Schluß vom Zornschen Lemma auf die Wohlordnung der Mächtigkeiten ist, von allem Vorangegangenen unabhängig, erst 1954 HÖNIG gelungen[13]). Trotz der Eleganz und Kürze des Hönigschen Beweises ist vielleicht auch die folgende, ZORNs Gedanken aufgreifende Herleitung immer noch von Interesse; man wird diesen neuen Beweis vielleicht als, bei gleicher Kürze, nicht minder natürlich ansehen.

Wir betrachten also eine beliebige nichtleere (mindestens ein Glied enthaltende) Mengenfamilie $(E_i)_{i\in I}$. Es gibt wenigstens einen Index $i\in I$, wir dürfen also einen Index $j\in I$ ein für allemal auszeichnen. T sei das cartesische Produkt[14]) der sämtlichen wie oben induktiv geordneten Mengen $\mathsf{T}(E_j, E_i)$ $(i\in I)$. Dann ist auch T induktiv geordnet, und das Supremum einer Kette $K\subseteqq T$ berechnet sich koordinatenweise. Jedem Element des Produkts — einer gewissen Familie $\varphi=(\varphi_i)_{i\in I}$ mit $\varphi_i\in\mathsf{T}(E_j, E_i)$ — entspricht eindeutig die Familie der Definitionsbereiche $D(\varphi_i)$ wie die der Wertevorräte $W(\varphi_i)$. Die Teilmenge S von T bestehe nun aus denjenigen Familien (φ_i), bei denen die Familie der Definitionsbereiche $D(\varphi_i)$ konstant ist, also alle $D(\varphi_i)$ gleich ein und derselben Teilmenge von E_j sind; man kann diese eindeutig bestimmte Teilmenge von E_j dann auch als den Definitionsbereich $D(\varphi)$ der Familie $\varphi=(\varphi_i)$ bezeichnen. Offenbar ist auch die Menge $S\subseteqq T$ induktiv: das in T genommene Supremum einer Kette $K\subseteqq S$ liegt immer noch in S, ist also obere Schranke, ja obere Grenze von K in S. Nach dem Maximalprinzip besitzt also die Menge S ein maximales Element $\mu=(\mu_i)$. Dann muß aber einer der beiden folgenden Fälle eintreten: es ist 1. $D(\mu)=E_j$ und damit μ_i eine eineindeutige Abbildung von E_j in E_i $(i\in I)$, E_j also die im Sinne der Mächtigkeit kleinste unter den Mengen E_i; oder 2. für wenigstens einen Index $k\in I$ gilt $W(\mu_k)=E_k$, wodurch die Hintereinanderausführung $\mu_i\cdot\mu_k^{-1}$ der Abbildungen μ_k^{-1} und μ_i eine eineindeutige Abbildung von E_k in E_i, E_k also die kleinste unter den Mengen E_i wird. Wäre weder das eine noch das andere der Fall, wäre also $D(\mu)\subset E_j$ und (für jedes $i\in I$) zugleich $W(\mu_i)\subset E_i$, so ließe sich, unter abermaliger Benutzung des Maximalprinzips in der speziellen Form des Auswahlaxioms, eine Familie $\nu\in S$ mit $\nu>\mu$ finden, wider die Maximaleigenschaft von μ.

---

[12]) FARAH [5].

[13]) HÖNIG [7].

[14]) Auch direktes oder (BIRKHOFF) kardinales Produkt: Menge der Auswahlfunktionen in koordinatenweiser Anordnung; s. BOURBAKI [3] p. 7f.

Literatur

[1] BACHMANN, H.: Transfinite Zahlen. Berlin 1955. — [2] BOURBAKI, N.: Théorie des ensembles (fascicule de résultats). Actual. Sci. Industr., 1. Aufl. **846** (1939); 2. Aufl. **1141** (1951). — [3] BOURBAKI, N.: Ensembles ordonnés. Cardinaux. Nombres entiers. Actual. Sci. Industr. **1243** (1956). — [4] FARAH, E.: Über die Totalordnung der Menge der Mächtigkeiten der Teilmengen einer gegebenen Menge. (Portugiesisch.) Bol. Soc. Mat. São Paulo **5**, 59—61 (1952). — [5] FARAH, E.: Sur le bon ordre de l'ensemble des puissances des parties d'un ensemble donné. Summa Brasil. Math. **3**, 37—42 (1953). — [6] FRAENKEL, A. A.: Abstract set theory. Amsterdam 1953. — [7] HÖNIG, CH. S.: Proof of the well-ordering of cardinal numbers. Proc. Amer. Math. Soc. **5**, 312 (1954). — [8] SEKI, S.: On transfinite inferences. Comment. Math. Univ. St. Pauli **4**, 43—45 (1955). — [9] ZORN, M.: Idempotency of infinite cardinals. Univ. Calif. Publ. Math. **2**, 9—12 (1944).

Köln, Mathematisches Seminar der Universität

(Eingegangen am 17. November 1956)

BASU, S. K.
Math. Zeitschr. Bd. 67, S. 303—309 (1957)

On comparison of the total strength of some Hausdorff methods

By
S. K. BASU

Introduction. The problem of total relative strength (for this definition see [2], p. 51, footnote) of some Hausdorff methods (H-methods) or matrices defining such methods (for properties of Hausdorff matrices see [6], Chapter XI or [7]) has already been considered by the author in some previous papers ([1], [3], [4]). The present note deals with the question of total relative strength of some more matrices of this kind (these matrices are listed in § 1 below).

1. The Hausdorff transformation of a *real* sequence $\{s_n\}$ is defined by

$$\sigma_n = \sum_{k=0}^{n} \binom{n}{k} \Delta^{n-k} \mu_k \cdot s_k \qquad (n = 0, 1, 2, \ldots), \tag{1}$$

where $\{\mu_n\}$ is a given factor sequence and

$$\Delta^0 \mu_k = \mu_k, \quad \Delta^n \mu_k = \Delta^{n-1} \mu_k - \Delta^{n-1} \mu_{k+1} \qquad (n \geq 1).$$

If H denotes the matrix transformation (1), we write $H \sim \mu_n$. The following are some of the special cases of H-methods with which we shall be concerned here:

$$\Gamma_\beta^\alpha \sim \left(\frac{\beta}{n+\beta}\right)^\alpha \qquad (\alpha, \beta > 0).$$

In particular, $\Gamma_1^\alpha = H^\alpha$, the Hölder matrix; and $\Gamma_\beta^1 = \Gamma_\beta \approx \Gamma_1 = \Gamma_1^1 = H^1$, the arithmetic mean matrix (the notation $\approx$ stands for the equivalence of the two matrices. For this definition see [6], p. 66).

Exponential matrix: $J^\beta \sim e^{-\beta n^{\frac{1}{2}}} \qquad (\beta > 0)$.

These matrices are all totally regular (for the definition of total regularity see [6], p. 53 or [1], p. 450 and for the conditions of total regularity of H-methods see [1], p. 451). The total regularity of H^α is well known and the same for Γ_β^α follows immediately. The total regularity of J^β is a direct consequence of the total monotony of the sequence $e^{-\beta n^{\frac{1}{2}}}$ (see [7], p. 97 or [6], p. 269).

It is well known that

$$\Gamma_\beta^\alpha \approx H^\alpha \qquad ([6], \text{p. } 265).$$

Regarding Γ_β^α the following theorem will be proved.

Theorem 1. (i) *The total strength of Γ_β^α increases along with α, β remaining constant.*

(ii) *The total strength of Γ_β^α increases with decreasing β, α remaining constant.*

Proof. (i) For the regular matrix $\Gamma_\beta^\alpha/\Gamma_\beta^{\alpha'}$ $(\alpha>\alpha'>0)$ (the regularity of this matrix follows immediately from the properties of Γ_β^α (see [7], p. 88)) the moment function $\mu(t)$ is given by

$$\mu(t)=\left(\frac{\beta}{t+\beta}\right)^\alpha\Big/\left(\frac{\beta}{t+\beta}\right)^{\alpha'}=\left(\frac{\beta}{t+\beta}\right)^{\alpha-\alpha'}.$$

Since

$$\Gamma_\beta^{\alpha-\alpha'}\sim\left(\frac{\beta}{t+\beta}\right)^{\alpha-\alpha'}\qquad(\alpha>\alpha'),$$

it follows that $\mu(t)$ is totally monotone and consequently, Γ_β^α is totally stronger than $\Gamma^{\alpha'}$.

(ii) Again, for $\Gamma_\beta^\alpha/\Gamma_{\beta'}^\alpha$ $(\beta'>\beta)$,

$$\mu(t)=(\beta/\beta')^\alpha\cdot\left(\frac{t+\beta'}{t+\beta}\right)^\alpha=e^{\varrho(t)},\quad\text{say.}$$

Then

$$\varrho(t)=\alpha\log\beta/\beta'+\alpha\{\log(t+\beta')-\log(t+\beta)\};$$

and

$$(-1)^k\varrho^{(k)}(t)=\alpha(k-1)!\left\{\frac{1}{(t+\beta)^k}-\frac{1}{(t+\beta')^k}\right\}>0\quad(t>0)\qquad(k=1,2,3,\ldots).$$

Consequently, as $\varrho(0)=\varrho(+0)$, $\mu(t)$ is totally monotone for $t\geq 0$ when $\alpha,\beta'-\beta>0$ (see [7], p. 95) and the result follows.

From Theorem 1, with the aid of a known lemma ([1], p. 452), we have at once the following results.

Theorem 2. *If $\alpha>0$,*

(i) *for $0<\beta<1$, $\Gamma_\beta^\alpha\ t\cdot s\cdot H^\alpha$ but $H^\alpha\ n\cdot t\cdot s\cdot\Gamma_\beta^\alpha$;*

(ii) *for $\beta>1$, $H^\alpha\ t\cdot s\cdot\Gamma_\beta^\alpha$ but $\Gamma_\beta^\alpha\ n\cdot t\cdot s\cdot H^\alpha$.*

(For the notations used here see [3], p. 390.)

Theorem 3.

(i) *For $\alpha>1$, $\Gamma_\beta^\alpha\ t\cdot s\cdot\Gamma_\beta$.* (ii) *When $0<\alpha<1$, $\Gamma_\beta\ t\cdot s\cdot\Gamma_\beta^\alpha$.*

(iii) *For $0<\beta<1$, $\Gamma_\beta\ t\cdot s\cdot H^\beta$.* (iv) *When $\beta>1$, $H^\beta\ t\cdot s\cdot\Gamma_\beta$.*

2. For the regular matrix J^α/J^β $(\alpha>\beta>0)$ the moment function

$$\mu(t)=e^{-(\alpha-\beta)t^{\frac{1}{\beta}}}$$

is totally monotone. Hence

$$J^\alpha\ t\cdot s\cdot J^\beta\qquad(\alpha>\beta>0);$$

so that the total strength of J^β increases with β.

It is well known that for $\alpha,\beta>0$, J^β is stronger than H^α (see [6], p. 270, Theorem 214). It is natural to inquire if J^β is also totally stronger than H^α. In this connection, the following theorem will be proved.

Theorem 4.

(i) *For* $\alpha \geq \beta$, $J^\beta n \cdot t \cdot s \cdot H^\alpha$. (ii) *For* $\beta \geq \left(\frac{2\pi}{e}\right)^{\frac{1}{2}} \alpha$, $J^\beta t \cdot s \cdot H^\alpha$.

For the proof of the theorem we require the following lemma.

Lemma 1. *The sequence* $\{u_n\}$ *for which*

$$u_n = \frac{\Gamma(n)}{\Gamma(n-\frac{1}{2})} \cdot \frac{(2n-1)^{(2n-1)/2}}{(2n)^n} \qquad (n \geq 2)$$

is $\uparrow$ *and* $\to \frac{1}{\sqrt{2e}}$.

Proof. We have

$$\frac{u_n}{u_{n+1}} = \left(\frac{n+1}{n}\right)^{n+1} \left(\frac{2n-1}{2n+1}\right)^{(2n+1)/2} = d^{\frac{1}{2}}, \text{ say.}$$

Then $$d = \frac{n+1}{n} \left\{\frac{(n+1)(2n-1)}{n(2n+1)}\right\}^{2n+1} = \frac{1+\frac{1}{n}}{\left(1+\frac{1}{2n^2+n-1}\right)^{2n+1}}.$$

Since $\frac{2n+1}{2n^2+n-1} > \frac{1}{n}$,

$$\left(1+\frac{1}{2n^2+n-1}\right)^{2n+1} > 1 + \frac{2n+1}{2n^2+n-1} > 1+\frac{1}{n}.$$

Hence $d<1$; and consequently, $u_n/u_{n+1}<1$, so that $u_n \uparrow$.

Again,

$$u_n = \left(\frac{2n-1}{2n}\right)^{(2n-1)/2} \cdot \frac{\Gamma(n)}{\sqrt{2n}\,\Gamma\left(n-\frac{1}{2}\right)}$$

$$= \frac{1}{\left(1+\frac{1}{2n-1}\right)^{(2n-1)/2}} \cdot \frac{\Gamma(n)}{\sqrt{2n}\,\Gamma\left(n-\frac{1}{2}\right)} \to \frac{1}{\sqrt{e}} \cdot \frac{1}{\sqrt{2}} = \frac{1}{\sqrt{2e}},$$

since $\frac{\Gamma(n)}{\Gamma\left(n-\frac{1}{2}\right)} \sim n^{\frac{1}{2}}$. This proves the lemma.

Proof of Theorem 4. For the regular matrix J^β/H^α the moment function is

$$\mu(t) = (t+1)^\alpha e^{-\beta t^{\frac{1}{2}}} \qquad (\alpha, \beta > 0). \tag{2}$$

Hence for $t>0$,

$$\mu'(t) = \alpha(t+1)^{\alpha-1} e^{-\beta t^{\frac{1}{2}}} - \frac{\beta}{2}(t+1)^\alpha e^{-\beta t^{\frac{1}{2}}} t^{-\frac{1}{2}}.$$

Therefore,

$$\mu'(1) > 0, \quad \text{as} \quad \alpha > \beta.$$

Consequently, for $\alpha > \beta$, $\mu(t)$ is not totally monotone for $t \geq 0$; and this implies that

$$J^\beta n \cdot t \cdot s \cdot H^\alpha \quad \text{for} \quad \alpha > \beta.$$

Again, for $\alpha=\beta$ as $\mu'(1)=0$, $\mu(t)$ cannot be totally monotone for $t\geq 0$ (see [8], pp. 167—168) and the proof of (i) is complete.

For the proof of (ii) (here we have essentially followed the method indicated very briefly by HAUSDORFF ([7], p. 98)) we put

$$\mu(t)=(t+1)^{\alpha}e^{-\beta t^{\frac{1}{2}}}=e^{\varrho(t)}.$$

Then

$$\varrho(t)=\alpha\log(t+1)-\beta t^{\frac{1}{2}};$$

and

$$-\varrho'(t)=\frac{\beta}{2}t^{-\frac{1}{2}}-\frac{\alpha}{t+1}>0 \qquad \text{for } t>0, \text{ since } \beta>\alpha.$$

Again,

$$(-1)^n\varrho^{(n)}(t)=\beta\cdot\frac{1\cdot 3\cdot 5\ldots(2n-3)}{2^n}t^{-\frac{2n-1}{2}}-\alpha\frac{(n-1)!}{(t+1)^n} \qquad (n\geq 2).$$

Therefore, for $t>0$,

$$(-1)^n\varrho^{(n)}(t)\geq 0 \qquad (n=2,3,4,\ldots),$$

if

$$\beta\frac{(t+1)^n}{t^{(2n-1)/2}}\geq 2\alpha\sqrt{\pi}\frac{\Gamma(n)}{\Gamma(n-\frac{1}{2})} \qquad (n=2,3,4,\ldots;t>0). \tag{3}$$

On putting

$$f(t)=\frac{(t+1)^n}{t^{(2n-1)/2}} \qquad (t>0),$$

it is easily seen that for a given n,

$$\min f(t)=\frac{(2n)^n}{(2n-1)^{(2n-1)/2}}.$$

Hence (3) certainly holds for $t>0$ if

$$\frac{\beta}{2\sqrt{\pi}\alpha}\geq\frac{\Gamma(n)}{\Gamma(n-\frac{1}{2})}\cdot\frac{(2n-1)^{(2n-1)/2}}{(2n)^n} \qquad (n\geq 2)$$

or, if

$$\frac{\beta}{2\sqrt{\pi}\alpha}\geq\max(u_2,u_3,u_4,\ldots), \tag{4}$$

where

$$u_n=\frac{\Gamma(n)}{\Gamma(n-\frac{1}{2})}\cdot\frac{(2n-1)^{(2n-1)/2}}{(2n)^n}.$$

Now it follows from Lemma 1 that

$$\text{l.u.b. of } u_n=\frac{1}{\sqrt{2e}}.$$

Accordingly, (4) will certainly be satisfied if

$$\frac{\beta}{2\sqrt{\pi}\alpha}\geq\frac{1}{\sqrt{2e}}, \qquad \text{that is, if} \qquad \beta\geq\left(\frac{2\pi}{e}\right)^{\frac{1}{2}}\alpha.$$

Since $\varrho(0)=\varrho(+0)$, it, therefore, follows that $\mu(t)$ is totally monotone for $t\geq 0$ when $\beta\geq\left(\frac{\pi e}{2}\right)^{\frac{1}{2}}\alpha$. This completes the proof of (ii).

3. After the proof of Theorem 4 the question naturally arises as to whether or not

$$J^{\beta}t\cdot s\cdot H^{\alpha}$$

for $\alpha<\beta<\left(\frac{2\pi}{e}\right)^{\frac{1}{2}}\alpha$. While we are unable to answer this question completely, the next theorem gives a partial answer.

Theorem 5. *For* $\alpha<\beta<\frac{3\sqrt{3}}{4}\alpha$, $J^{\beta}n\cdot t\cdot s\cdot H^{\alpha}$.

We require the following lemma.

Lemma 2. *If* $f(t)$ *is totally monotone for* $t\geq 0$ *and not identically zero, then it is logarithmically convex for* $t\geq 0$ (see [8], p. 167).

Proof of the Theorem. We have from (2),

$$\log\mu(t)=\alpha\log(t+1)-\beta t^{\frac{1}{2}}.$$

Therefore,

$$\frac{d}{dt}\log\mu(t)=\frac{\alpha}{t+1}-\frac{\beta}{2}t^{-\frac{1}{2}},$$

and

$$\frac{d^2}{dt^2}\log\mu(t)=-\frac{\alpha}{(t+1)^2}+\frac{\beta}{4}t^{-\frac{3}{2}}.$$

Hence

$$\frac{d^2}{dt^2}\log\mu(t)<0 \text{ for some } t>0,$$

if there is some $t>0$ for which

$$\frac{\alpha t^{\frac{3}{2}}}{(t+1)^2}>\frac{\beta}{4}. \tag{5}$$

Now, for $t>0$ the function

$$\varphi(t)=\frac{t^{\frac{3}{2}}}{(t+1)^2}$$

is maximum at $t=3$. Then (5) holds with $t=3$, if

$$\frac{3\sqrt{3}}{16}\alpha>\frac{\beta}{4}, \quad \text{that is} \quad \beta<\frac{3\sqrt{3}}{4}\alpha.$$

Thus for $\alpha<\beta<\frac{3\sqrt{3}}{4}\alpha$, by Lemma 2, $\mu(t)$ cannot be totally monotone and consequently,

$$J^{\beta}n\cdot t\cdot s\cdot H^{\alpha}.$$

4. As the arithmetic mean method plays an important rôle in the theory of summability it may not be out of place to give here a complete treatment

for the comparison of the total strength of this method and the exponential method. This is contained in our last theorem.

Theorem 6. *For* $\beta \geq \frac{3}{2}$, $J^\beta t \cdot s \cdot H^1$;

for $\beta < \frac{3}{2}$, $J^\beta n \cdot t \cdot s \cdot H^1$.

Proof. We note that a sequence $\{\mu_n\}$ is totally monotone if and only if it has a representation

$$\mu_n = \int_0^1 u^n \, d\chi(u),$$

where the mass function $\chi(u) \geq 0$ is bounded and $\uparrow$ in $[0, 1]$ (see [7], pp. 102 to 103 or [1], p. 451). Suppose now that $\chi(u)$ is continuous in $[0, 1]$ and has in $(0, 1)$ a continuous derivative $\chi'(u) = \varphi(u)$, so that

$$\mu_n = \int_0^1 u^n \varphi(u) \, du.$$

Putting $u = e^{-t}$ we have, with $\Phi(t) = u\varphi(u) \geq 0$ for $t > 0$,

$$\mu_n = \int_0^\infty e^{-nt} \Phi(t) \, dt.$$

Let us now take $\mu_n = e^{-\beta n^{\frac{1}{2}}}$, so that

$$\Phi(t) = \frac{\beta}{2\sqrt{\pi}} e^{-\beta^2/(4t)} \cdot t^{-\frac{3}{2}} \quad (\beta, t > 0) \qquad \text{(see [7], p. 91)}.$$

For the regular matrix J^β/H^1 the moment sequence $\{M_n\}$ is given by

$$\begin{aligned} M_n &= (n+1)\, e^{-\beta n^{\frac{1}{2}}} \\ &= \int_0^\infty [\Phi(t) + \Phi'(t)]\, e^{-nt} dt \qquad \text{([7], p. 90)} \\ &= \int_0^\infty e^{-nt} \Psi(t) \, dt, \end{aligned}$$

where $\Psi(t) = \Phi(t) + \Phi'(t)$ is continuous for $t > 0$. Hence $\{M_n\}$ is totally monotone if and only if

$$\Psi(t) \geq 0 \quad \text{for} \quad t > 0,$$

that is,

$$\frac{\beta}{2\sqrt{\pi}} e^{-\beta^2/(4t)} \cdot t^{-\frac{3}{2}} + \frac{\beta}{2\sqrt{\pi}} e^{-\beta^2/(4t)} \cdot \frac{\beta^2}{4t^2} t^{-\frac{3}{2}} - \frac{\beta}{2\sqrt{\pi}} e^{-\beta^2/(4t)} \cdot \frac{3}{2} t^{-\frac{5}{2}} \geq 0,$$

or

$$4t^2 - 6t + \beta^2 \geq 0 \quad \text{for} \quad t > 0; \tag{6}$$

and this will be the case if and only if $\beta \geq \frac{3}{2}$. The result now follows.

For $\alpha = k$, a positive integer we have from (2),

$$\nu(t) \equiv (t+1)^k e^{-\beta t^{\frac{1}{2}}}$$
$$= \left[(t+1)e^{-\beta k \cdot t^{\frac{1}{2}}}\right]^k.$$

It follows from the proof of Theorem 6 that the moment function $(t+1)e^{-\beta t^{\frac{1}{2}}}$ corresponding to $\{M_n\}$ is totally monotone for $\beta \geq \frac{3}{2}$ (see [1], Lemma 1 or [7], pp. 102—103). Hence $\nu(t)$ is totally monotone for $\beta/k \geq \frac{3}{2}$ (see [3], Lemma 4) and this implies that

$$J^k t \cdot s \cdot H^\beta \quad \text{for} \quad \beta \geq \tfrac{3}{2} \cdot k.$$

We have thus a stronger result than Theorem 4 (ii) when α is a positive integer. The above simple method, however, does not seem to apply to the case for non-integral values of α.

In conclusion, we would like to remark that the general question as to whether or not $J^\beta t \cdot s \cdot H^\alpha$ for $3\sqrt{3}/4\alpha \leq \beta < \left(\frac{2\pi}{e}\right)^{\frac{1}{2}}\alpha$ still remains open, though in some sense this gap is covered by Theorem 6.

References

[1] Basu, S. K.: On the total relative strength of the Hölder and Cesàro methods. Proc. London Math. Soc. (2) **50**, 447—462 (1948). — [2] Basu, S. K.: On the total relative strength of the Riesz and Cesàro methods. J. London Math. Soc. **24**, 51—59 (1949). — [3] Basu, S. K.: On the total relative strength of some Hausdorff methods equivalent to identity. Amer. J. Math. **76**, 389—398 (1954). — [4] Basu, S. K.: On hypergeometric summability involving infinite limits. Proc. Amer. Math. Soc. **5**, 226—238 (1954). — [5] Garabedian, H. L., and H. S. Wall: Hausdorff methods of summation and continued fractions. Trans. Amer. Math. Soc. **48**, 185—207 (1940). — [6] Hardy, G. H.: Divergent Series. Oxford 1949. — [7] Hausdorff, F.: Summationsmethoden und Momentfolgen. I. Math. Z. **9**, 74—109 (1921). — [8] Widder, D. V.: The Laplace transform. Princeton 1946.

The Johns Hopkins University and H. M. College, West Bengal

(Eingegangen am 14. August 1956)

RAJAGOPAL, C. T.
Math. Zeitschr. Bd. 67, S. 310—319 (1957)

On a theorem of Frobenius and Knopp for Abel summability

By
C. T. RAJAGOPAL

Introduction

It is a classical result that sequences summable either (C, k) or (H, k), for some $k>0$, are also summable (A), i.e. summable by the Abel power-series transform, to the same sum[1]). In the case $k=1$, the result was proved by FROBENIUS [1]. In the same case, $k=1$, KNOPP [6] proved an analogous result for functions which states that $s(u)$, summable by the integral transform called its (C, k) or (H, k) mean, is summable also by the Laplace-Abel integral transform to the same sum. KNOPP's result is true for $0<k<1$, being, in fact, contained in his result for $k=1$; but it is not necessarily true for $k>1$[2]). However, there is an *indirect* extension of KNOPP's result to all $k=r_0+1>1$, provided by Theorem 1′ of this note when its $f(t)\equiv 1$. Theorem 1 preceding Theorem 1′ is an exact analogue of Theorem 1′ for a sequence s_n[3]), and its case $r_0=0$ is a result recently proved by POSTNIKOV [8].

Theorems 2 and 2′ are limiting forms of Theorems 1 and 1′, where the hypothesis of summability (σ, r_0+1), i.e. either (C, r_0+1) or (H, r_0+1), restricting s_n in Theorem 1 and $s(u)$ in Theorem 1′, is replaced by the hypothesis of summability (C, ∞) suitably defined. Theorems 2 and 2′ partly depend on two well-known theorems of KARAMATA, Propositions III and III′ of this note, and so, by implication, on the idea of WEIERSTRASS' polynomial approximations. The same idea is employed by POSTNIKOV, but it has been avoided, with a gain in simplicity, in the proofs of Theorems 1 and 1′ which generalize his theorem.

1. Notation

Given any sequence s_n $(n=0, 1, 2, \ldots)$, we denote by $C_n^k(s_\nu)\equiv C_n^k$ and $H_n^k(s_\nu)\equiv H_n^k$ its (C, k) and (H, k) means respectively, for all $k>0$ ([2], pp. 96, 250–1); and we define (C, k) and (H, k) summabilities of s_n, each to l (finite), by

$$(1)\qquad C_n^k \equiv S_n^k \Big/ \binom{n+k}{n} = \sum_{\nu=0}^{n} \binom{\nu+k-1}{\nu} s_{n-\nu} \Big/ \binom{n+k}{n} \to l \quad (n\to\infty),$$

---

[1]) The notation, as far as practicable, conforms to that of HARDY [2].

[2]) This is shown by KNOPP's example of $s(u)$ summable (C, k), $k>1$, and so equivalently summable (H, k), for which the Laplace-Abel transform does not exist.

[3]) Corresponding results, relations and definitions for a sequence s_n and a function $s(u)$ are given the same number throughout this note, but the number in the case of s_n is unaccented (e.g. 1, I, ...) while that in the case of $s(u)$ is accented (e.g. 1′, I′, ...).

$$(2) \qquad H_n^k \equiv \sum_{\nu=0}^{n} (-1)^\nu \binom{n}{\nu} \Delta^\nu H_0^k = \sum_{\nu=0}^{n} (-1)^\nu \binom{n}{\nu} \frac{\Delta^\nu s_0}{(\nu+1)^k} \to l \quad (n \to \infty),$$

respectively. Summability (A) of s_n to l we define as usual by

$$(3) \qquad A(y; s_\nu) = y \sum_{\nu=0}^{\infty} s_\nu e^{-y\nu} \to l \quad \text{as} \quad y \to +0,$$

or equivalently, by

$$(3a) \qquad A(x; s_\nu) = (1-x) \sum_{\nu=0}^{\infty} s_\nu x^\nu \to l \quad \text{as} \quad x \to 1-0,$$

where $A(y; s_\nu)$ or $A(x; s_\nu)$ is a convergent series for $y>0$ or $0<x<1$, called the Abel or (A) power-series transform of the sequence s_n.

Next let $s(u)$ denote, as everywhere in this note, unless otherwise stated, a function bounded and Lebesgue-integrable in every finite interval of $u \geq 0$. Then, for any $k>0$ and $u>0$, we define as usual (e.g. [5], p. 232) $C^k(u; s(x)) \equiv C^k(u)$ and $H^k(u; s(x)) \equiv H^k(u)$, the (C, k) and (H, k) means respectively of $s(u)$, and summabilities (C, k) and (H, k) of $s(u)$, each to l, by

$$(1') \qquad C^k(u) \equiv \frac{\Gamma(k+1)\, S^k(u)}{u^k} = \frac{k}{u^k} \int_0^u (u-x)^{k-1} s(x)\, dx \to l \quad (u \to \infty),$$

$$(2') \qquad H^k(u) = \frac{1}{\Gamma(k)} \frac{1}{u} \int_0^u \left(\log(u/x)\right)^{k-1} s(x)\, dx \to l \quad (u \to \infty),$$

respectively. Summability (A) of $s(u)$ to l we define by

$$(3') \qquad A\left(y; s(u)\right) = y \int_0^\infty e^{-yu} s(u)\, du \to l \quad \text{as} \quad y \to +0,$$

where the integral is supposed to be convergent for $y>0$ in the Cauchy-Lebesgue sense and called the Laplace-Abel or (A) integral transform of $s(u)$.

Summability (C, ∞) of s_n [or $s(u)$] to l involves

$$(4 \text{ or } [4']) \qquad \limsup_{n \to \infty \text{ [or } u \to \infty]} C_n^r \text{ [or } C^r(u)] = \overline{C}_r, \qquad \liminf_{n \to \infty \text{ [or } u \to \infty]} C_n^r \text{ [or } C^r(u)] = \underline{C}_r,$$

and it is defined:

$$(5) \qquad \overline{C}_\infty \equiv \lim_{r \to \infty} \overline{C}_r = l, \quad \underline{C}_\infty \equiv \lim_{r \to \infty} \underline{C}_r = l,$$

where $\overline{C}_\infty, \underline{C}_\infty$ exist in any case, as finite numbers or not, on account of $\overline{C}_r$ in (4 or [4']) being a monotonic decreasing function of r and $\underline{C}_r$ in (4 or [4']) a monotonic increasing function of r.

Working out a notation already indicated, *we write*

$$(6) \qquad \sigma_n^k = \textit{either } C_n^k \textit{ of } (1) \textit{ or } H_n^k \textit{ of } (2),$$

$$(6') \qquad \sigma^k(u) = \textit{either } C^k(u) \textit{ of } (1') \textit{ or } H^k(u) \textit{ of } (2'),$$

with the understanding that σ, when it appears in any theorem, is to be replaced throughout by each one of C, H in turn. Further, writing $\sigma_n^0 = s_n$ and $\sigma^0(u) = s(u)$, we define summability (σ, k) to l for all $k \geq 0$ as $\sigma_n^k \to l$ $(n \to \infty)$ or $\sigma^k(u) \to l$ $(u \to \infty)$ according as we are thinking of sequences s_n or functions $s(u)$.

2. Auxiliary propositions

Of the results in this section, Propositions I and I′ give the essential content of Theorems 1 and 1′ while Propositions II and II′ taken along with Propositions III and III′ lead to Theorems 2 and 2′.

Proposition I. *If a_n $(n = 0, 1, 2, \ldots)$ is a sequence summable $(C, 1)$ to l and if*

$$\text{(7)} \qquad \begin{cases} \psi(u) \text{ is positive monotonic decreasing for } u > 0, \\ \int_0^\infty \psi(u)\, du < \infty, \end{cases}$$

then

$$\text{(8–i)} \qquad y \sum_0^\infty a_n \psi(yn) \quad \text{converges for } y > 0,$$

$$\text{(8–ii)} \qquad y \sum_0^\infty a_n \psi(yn) \quad \text{tends to} \quad l \int_0^\infty \psi(u)\, du \quad \text{as } y \to +0.$$

Proof. Writing

$$\Psi(y, M, N) = y \sum_M^N a_n \psi(yn), \quad y > 0, \quad N \geq M + 1,$$

where M is a fixed positive integer, we shall prove (8–i) and (8–ii) in the form:

$$\text{(9)} \qquad \lim_{N \to \infty} \Psi(y, M, N) \text{ exists}, \qquad \lim_{y \to +0} \lim_{N \to \infty} \Psi(y, M, N) = l \int_0^\infty \psi(u)\, du.$$

If we set $b_n = a_0 + a_1 + \cdots + a_n$ and use our hypothesis on a_n to choose M so that, for $n \geq M - 1$, $b_n = (1 + \varepsilon_n) nl$ where $|\varepsilon_n| < \varepsilon$ and $\varepsilon > 0$ is any small number, we get

$$\text{(10)} \qquad \begin{aligned} \Psi(y, M, N) &= y \sum_M^N (b_n - b_{n-1})\, \psi(yn) \\ &= y \sum_M^{N-1} b_n [\psi(yn) - \psi(y\overline{n+1})] - y\, b_{M-1} \psi(yM) + y\, b_N \psi(yN) \\ &= l\, y \sum_M^{N-1} (1 + \varepsilon_n)\, n\, [\psi(yn) - \psi(y\overline{n+1})] - \\ &\quad - l\, y (1 + \varepsilon_{M-1})(M - 1)\, \psi(yM) + l\, y (1 + \varepsilon_N) N \psi(yN) \\ &= l\, y \sum_M^N \psi(yn) + R(y, M, N), \end{aligned}$$

where $R(y, M, N)$ is easily shown to satisfy the condition

$$|R(y, M, N)| \leq l\, y\, \varepsilon \sum_{M}^{N} \psi(y\, n) + 2l\, y\, \varepsilon\, (M-1)\, \psi(y\, M).$$

Now it is clear that

$$y \int_{M}^{N+1} \psi(y\, u)\, du \leq y \sum_{M}^{N} \psi(y\, n) \leq y \int_{M-1}^{N} \psi(y\, u)\, du \leq y \int_{0}^{\infty} \psi(y\, u)\, du.$$

Hence

$$\lim_{N\to\infty} y \sum_{M}^{N} \psi(y\, n) \text{ exists}, \quad \lim_{y\to+0} \lim_{N\to\infty} y \sum_{M}^{N} \psi(y\, n) = \int_{0}^{\infty} \psi(u)\, du. \tag{11}$$

Using (11) in (10), we obtain (9) as we wished to. For, $u\psi(u)\to 0$ as $u\to\infty$ by (7) and so $y(M-1)\psi(yM) < yM\psi(yM) < k$ independent of y and M; consequently

$$\limsup_{N\to\infty} |R(y, M, N)| \leq l\, \varepsilon \int_{0}^{\infty} \psi(u)\, du + 2l\, k\, \varepsilon,$$

which means, ε being arbitrary,

$$\lim_{y\to+0} \lim_{N\to\infty} R(y, M, N) = 0.$$

Proposition I'. *If $b(u)$ is a continuous function of bounded variation in every finite interval of $u \geq 0$,*

$$b(0) = 0, \quad b(u) \sim l\, u \quad \text{as} \quad u \to \infty,$$

and if $\psi(u)$ is as in (7), then

(8'–i) $\int_{0}^{\infty} d\, b(u) \cdot \psi(y\, u)$, $y > 0$, *exists as a Cauchy-Lebesgue-Stieltjes integral, i.e. exists as the limit when $x \to \infty$ of the corresponding Lebesgue-Stieltjes integral from 0 to x, and*

(8'–ii) $$y \int_{0}^{\infty} d\, b(u) \cdot \psi(y\, u) \to l \int_{0}^{\infty} \psi(u)\, du \quad \text{as} \quad y \to +0.$$

Proof. Since we may define $b(u) = 0$ for $u < 0$ and $\psi(u) = \psi(+0)$ for $u \leq 0$, we get by an integration by parts

$$\int_{0}^{x} d\, b(u) \cdot \psi(y\, u) = b(x)\, \psi(y\, x + 0) - \int_{0}^{x} b(u) \cdot d\psi(y\, u).$$

When $x \to \infty$ (y remaining fixed), $b(x)\psi(yx+0) \sim x\psi(yx+0) \to 0$ since $u\,\psi(u) \to 0$ as $u \to \infty$. Hence it is enough to prove that, for fixed $U > 0$,

(9') $$\begin{cases} \lim_{x\to\infty} -\, y \int_{U}^{x} b(u) \cdot d\psi(y\, u) \text{ exists for } y > 0, \\ \quad \text{tending to } l \int_{0}^{\infty} \psi(u)\, du \text{ as } y \to +0. \end{cases}$$

We prove this like the corresponding step in Proposition I; i.e. we choose U so that, for $u \geq U$, $b(u) = (1+\varepsilon_u)ul$, $|\varepsilon_u| < \varepsilon$, and write

$$(10') \quad -y\int_U^x b(u)\cdot d\psi(u) = -ly\int_U^x u\,d\psi(yu) - ly\int_U^x \varepsilon_u u\,d\psi(yu) = I_1 + I_2 \qquad (x > U),$$

where it is easy to see, by an integration by parts, that

$$(11') \quad \lim_{x\to\infty} I_1 \text{ exists}, \quad \lim_{y\to+0}\lim_{x\to\infty} I_1 = l\int_0^\infty \psi(u)\,du.$$

Since $|I_2| < \varepsilon I_1$, (10') and (11') together lead to our conclusion in the form (9').

Proposition II. *For a sequence s_n $(n = 0, 1, 2, \ldots)$, the hypothesis* (5) *of*

summability (C, ∞) *to* l

leads to the conclusion of

summability (A) *to* l *of the* (C, r) *mean for* $r > 0$.

Proof. The proposition states that the summability in the conclusion, which may be called summability $(A)\,(C, r)$, $r > 0$, includes summability (C, ∞); and it can be symbolically written:

$$(A)\,(C, r) > (C, \infty), \quad r > 0.$$

In this form the proposition is obvious from the relations

$$(A) > (C, \infty), \quad (A)\,(C, r) > (A), \quad r > 0,$$

of which the first is well known (e.g. [7], p. 199) and the second is a theorem due to Szász ([12], p. 1234).

Proposition II'. *For $s(u)$ defined as in* § 1, *the hypothesis* (5) *of*

summability (C, ∞) *to* l

implies the conclusion of

summability (A) *to* l *of the* (C, r) *mean for* $r \geq r_0 > 0$,

where r_0 is such that $\underline{C}_r, \overline{C}_r$ are bounded for $r \geq r_0$ (its existence following from our hypothesis).

Proof. We have to show that, for each $r \geq r_0$,

$$(12) \quad A\left(y; C^r(u)\right) \equiv y\int_0^\infty e^{-yu} C^r(u)\,du \to l \text{ as } y \to +0,$$

where the integral is clearly absolutely convergent for $y > 0$. (12) for $r = r_0$ is proved in three stages numbered below (13), (14), (15). (12) for any given $r > r_0$ is proved in the same way.

$$(13) \quad A^{r_0}(y) \equiv y^{r_0+1}\int_0^\infty e^{-yu} S^{r_0}(u)\,du = y^{r_0+k+1}\int_0^\infty e^{-yu} S^{r_0+k}(u)\,du \equiv A^{r_0+k}(y) \qquad (y > 0, k > 0);$$

(14) $$A^{r_0}(y) \to l \quad \text{as} \quad y \to +0;$$

(15) $$y\int_0^\infty e^{-yu} C^{r_0}(u)\,du = \int_y^\infty A^{r_0}(u)\cdot d\left(1-\frac{y}{u}\right)^{r_0} \to l \quad \text{as} \quad y \to +0.$$

To prove (13) we use a well-known formula which connects $S^{r_0+k}(u)$ and $S^{r_0}(u)$, obtaining first

$$A^{r_0+k}(y) = \frac{y^{r_0+k+1}}{\Gamma(k)}\int_0^\infty e^{-yu}\,du\int_0^u (u-x)^{k-1} S^{r_0}(x)\,dx$$

$$= \frac{y^{r_0+k+1}}{\Gamma(k)}\int_0^\infty S^{r_0}(x)\,dx\int_x^\infty e^{-yu}(u-x)^{k-1}\,du,$$

where the inversion of integrations is justified by FUBINI's theorem. We have then only to write

$$A^{r_0+k}(y) = \frac{y^{r_0+k+1}}{\Gamma(k)}\int_0^\infty e^{-yx} S^{r_0}(x)\,dx\int_0^\infty e^{-yt} t^{k-1}\,dt \qquad (t=u-x)$$

$$= \frac{y^{r_0+1}}{\Gamma(k)}\int_0^\infty e^{-yx} S^{r_0}(x)\,dx\,\Gamma(k) \equiv A^{r_0}(y).$$

To prove (14) we use in (13) the expression for $S^{r_0+k}(u)$ in terms of $C^{r_0+k}(u)$ given by (1'), and get as a first step

$$\underline{C}_{r_0+k} \le \lim_{y\to+0}{}^{\sup}_{\inf} A^{r_0}(y) \le \bar{C}_{r_0+k},$$

where $\underline{C}_{r_0+k}$, $\bar{C}_{r_0+k}$ are defined as in (4'). (14) follows from this step when we let $k\to\infty$ and use our hypothesis $\bar{C}_\infty = \underline{C}_\infty = l$.

Finally, to prove (15) we start with the following observation which is a result of the definition of $A^{r_0}(u)$ in (13):

$$\int_y^\infty A^{r_0}(u)\cdot d\left(1-\frac{y}{u}\right)^{r_0} = r_0\, y\int_y^\infty \frac{A^{r_0}(u)}{u^{r_0+1}}(u-y)^{r_0-1}\,du$$

$$= r_0\, y\int_y^\infty (u-y)^{r_0-1}\,du\int_0^\infty e^{-ux} S^{r_0}(x)\,dx.$$

Appealing to FUBINI's theorem and inverting the order of integrations, we get

$$\int_y^\infty A^{r_0}(u)\cdot d\left(1-\frac{y}{u}\right)^{r_0} = r_0\, y\int_0^\infty S^{r_0}(x)\,dx\int_y^\infty e^{-ux}(u-y)^{r_0-1}\,du$$

$$= r_0\, y\int_0^\infty e^{-yx} S^{r_0}(x)\,dx\int_0^\infty e^{-tx} t^{r_0-1}\,dt \qquad (t=u-y)$$

$$= y\int_0^\infty e^{-yx}\,\frac{S^{r_0}(x)\,\Gamma(r_0+1)}{x^{r_0}}\,dx = y\int_0^\infty e^{-yx} C^{r_0}(x)\,dx.$$

After this we can establish the limit relation in (15) by writing

$$(16) \qquad \int_{y}^{\infty} A^{r_0}(u)\cdot d\left(1-\frac{y}{u}\right)^{r_0} = \int_{1}^{0} A(y/t)\cdot d(1-t)^{r_0} \qquad (y/u=t)$$

where $|A^{r_0}(y/t)| \equiv |A^{r_0}(u)| < K\Gamma(r_0+1)$, K being a constant independent of y and t $(y>0,\ 0<t<1)$ by our assumption that $|S^{r_0}(u)| < Ku^{r_0}$ for $0<u<\infty$. Since, for each t such that $0<t<1$, $A^{r_0}(y/t)\to l$ as $y\to+0$ by (14), (16) yields, when we appeal to the Lebesgue theorem of dominated convergence:

$$\int_{y}^{\infty} A^{r_0}(u)\cdot d\left(1-\frac{y}{u}\right)^{r_0} \to \int_{1}^{0} l\, d(1-t)^{r_0} = l.$$

This concludes the proof.

Remarks on Propositions II, II'. (i) Proposition II is proved by me elsewhere ([9], Lemma II) in the incomplete form (14) by a circuitous method. (ii) The distinction between the (A) transforms of s_n and $s(u)$, brought to light in KNOPP's theorem mentioned in the Introduction, is reflected in the conclusions of Propositions II, II', and it can be further elaborated as follows.

If $\bar{C}_\infty, \underline{C}_\infty$ are finite, then

for s_n, $\limsup_{x\to 1-0} A(x) = \bar{C}_\infty$, $\liminf_{x\to 1-0} A(x) = \underline{C}_\infty$ ([11], Theorem 1 (c)),

where $A(x)\equiv A(x; s_\nu)$ is defined by (3a), *while*

for $s(u)$, $\limsup_{y\to+0} A^r(y) = \bar{C}_\infty$, $\liminf_{y\to+0} A^r(y) = \underline{C}_\infty$, $r \gtrsim r_0 > 0$ ([10], Theorem 1_σ),

where $A^r(y)$ is defined as in (13).

Proposition III. *If s_n $(n=0, 1, 2, \ldots)$ is a sequence bounded below and summable (A) to l, and if $g(t)$ is Riemann-integrable over $0\le t\le 1$, then*

$$\lim_{x\to 1-0}(1-x)\sum_{n=0}^{\infty} s_n x^n g(x^n) = l\int_0^1 g(t)\,dt.$$

In particular, by a special choice of $g(t)$, we infer that s_n is summable $(C, 1)$ to l.

Proposition III'. *If $s(u)$ is bounded below, Riemann-integrable in every finite interval of $u\ge 0$, and summable (A) to l, and if $g(t)$ is Riemann-integrable over $0\le t\le 1$, then*

$$\lim_{y\to+0} y\int_0^{\infty} e^{-yu} g(e^{-yu})\, s(u)\,du = l\int_0^1 g(t)\,dt,$$

the integral existing in the Cauchy-Riemann sense for $y>0$.

A particular case of the conclusion, corresponding to a particular choice of $g(t)$, is that $s(u)$ is summable $(C, 1)$ to l.

KARAMATA has proved both Proposition III [3] and Proposition III' ([4], Hauptsatz 1 with $L(x)\equiv 1$, $\sigma=1$ and Satz 1 as a particular case).

3. Main theorems

Theorem 1. *If, for an* $r_0 \geq 0$, *the sequence* s_n $(n = 0, 1, 2, \ldots)$ *is summable* $(\sigma, r_0 + 1)$ *to* l, *i.e.*

(17) $\qquad \sigma_n^{r_0+1}$, *defined as in* (6), *tends to* l *as* $n \to \infty$,

then, for any function $f(t)$ *of bounded variation over* $0 \leq t \leq 1$ *and* $r \geq r_0$ [4]),

$$(18) \qquad y \sum_0^\infty e^{-yn} f(e^{-yn})\, \sigma_n^r \begin{cases} \text{converges for } y > 0, \\ \text{tends to } l \int_0^1 f(t)\, dt \text{ as } y \to +0. \end{cases}$$

Proof. It suffices to prove (18) for $r = r_0$ since afterwards we can change r_0 to $r_0 + \delta$, $\delta > 0$, in (17) and repeat our proof. Also we need prove (18) only for $f(t)$ which is positive monotonic increasing and bounded over $0 \leq t \leq 1$ since the given $f(t)$ can be expressed as the difference of two functions each with the property last stated. In other words, we can simplify the condition assumed on $f(t)$ so that $e^{-u} f(e^{-u})$ behaves exactly like the $\psi(u)$ of (7).

Now it is well known that, for $r_0 \geq 0$,

$$(19) \qquad \begin{cases} \dfrac{1}{n+1} \displaystyle\sum_{\nu=0}^{n} H_\nu^{r_0} = H_n^{r_0+1}, \\ \dfrac{1}{n+1} \displaystyle\sum_{\nu=0}^{n} C_\nu^{r_0} = \dfrac{1}{r_0+1} C_n^{r_0+1} + \dfrac{r_0}{r_0+1} \left(\dfrac{1}{n+1} \displaystyle\sum_{\nu=0}^{n} C_\nu^{r_0+1} \right) [5]). \end{cases}$$

Consequently (17) implies (in fact, is equivalent to) the summability $(C, 1)$ of $\sigma_n^{r_0}$ to l. Hence the conclusion (18) follows at once from Proposition I with

$$a_n = \sigma_n^{r_0}, \qquad \psi(u) = e^{-u} f(e^{-u}).$$

Corollary 1. *On the hypothesis* (17), *we have also the following conclusion more general than* (18): *for* $r \geq r_0$,

$$y \sum_0^\infty \psi(yn)\, f\{\varphi(yn)\}\, \sigma_n^r \begin{cases} \text{converges for } y > 0, \\ \text{tends to } l \int_0^1 f(t)\, dt \text{ as } y \to +0, \end{cases}$$

where $\psi(u)$ *is subject to* (7) *and* $\varphi(u)$ *is defined:*

$$\varphi(u) \equiv \int_u^\infty \psi(x)\, dx, \qquad \varphi(0) = 1.$$

To pass from the proof of Theorem 1 to that of Corollary 1, we have only to replace $e^{-u} f(e^{-u})$ in the former by $\psi(u) f\{\varphi(u)\}$.

---

[4]) For $r \geq r_0 + 1$, (18) and the similar conclusions of Theorem 1' and Corollaries 1, 1' are trivial.

[5]) This relation is known to be true for $\Re(r_0) > 0$ ([5], relation (4.1)).

The functional analogues of Theorem 1 and Corollary 1, stated below, are proved like them, but by using the functional analogue of (19) and Proposition I' instead of (19) and Proposition I.

Theorem 1'. *If, for an* $r_0 \geq 0$, $s(u)$ *defined as in* § 1 *is summable* (σ, r_0+1) *to* l, *i.e.*

$$\sigma^{r_0+1}(u), \text{ defined as in (6')}, \text{ tends to } l \text{ as } u \to \infty,$$

and if, $f(t)$ *is of bounded variation over* $0 \leq t \leq 1$, *then, for* $r \geq r_0$,

$$y \int_0^\infty e^{-yu} f(e^{-yu})\, \sigma^r(u)\, du \begin{cases} \text{exists as a Cauchy-Lebesgue integral if } y>0, \\ \text{tends to } l \int_0^1 f(t)\, dt \quad \text{as} \quad y \to +0. \end{cases}$$

Corollary 1'. *In Theorem* 1',

$y \int_0^\infty e^{-yu} f(e^{-yu})\, \sigma^r(u)\, du$ *can be replaced by* $y \int_0^\infty \psi(yu)\, f\{\varphi(yu)\}\, \sigma^r(u)\, du$

where ψ *and* φ *are defined as in Corollary* 1, *there being no other change.*

Theorem 2. *If* s_n $(n=0, 1, 2, \ldots)$ *is summable* (C, ∞) *to* l *according to definition* (5), *and if* $g(t)$ *is Riemann-integrable over* $0 \leq t \leq 1$, *then, for* $r \geq r_0 > 0$,

$$(1-x) \sum_0^\infty x^n g(x^n)\, C_n^r \begin{cases} \text{converges absolutely for } 0<x<1, \\ \text{tends to } l \int_0^1 g(t)\, dt \quad \text{as} \quad x \to 1-0. \end{cases}$$

The proof is like that of Theorem 2' and omitted.

Theorem 2'. *If* $s(u)$ *defined as in* § 1 *is summable* (C, ∞) *to* l *according to definition* (5), *and if* $g(t)$ *is Riemann-integrable over* $0 \leq t \leq 1$, *then, for* $r \geq r_0 > 0$,

$$(20) \quad y \int_0^\infty e^{-yu} g(e^{-yu})\, C^r(u)\, du \begin{cases} \text{is an absolutely convergent Riemann integral for } y>0, \\ \text{tends to } l \int_0^1 g(t)\, dt \quad \text{as} \quad y \to +0. \end{cases}$$

Proof. There is an $r_0 > 0$ such that, for $r \geq r_0$, (i) the (C, r) mean of $s(u)$ is bounded obviously, (ii) the (C, r) mean is summable (A) to l by Proposition II'. The first half of (20) follows at once from (i) and the second half from (i), (ii) together by an appeal to Proposition III' with $C^r(u)$ in place of $s(u)$.

Some special cases. Besides the special cases of the main theorems mentioned at the outset, viz. KNOPP's and POSTNIKOV's, the following may be noticed. In Corollaries 1 and 1', we can choose $\psi(u) = e^{-u}$ (Laplace-Abel kernel), getting Theorems 1 and 1', or else,

$$\text{either } \psi(u) = \frac{1}{(1+u)^{\varrho+1}}, \qquad \varrho > 0 \text{ (Stieltjes kernel)},$$

$$\text{or} \qquad \psi(u) = -\frac{d}{du}\left(\frac{u}{e^u - 1}\right) \qquad \text{(Lambert kernel)}.$$

Again, Theorems 2 and 2′, with the particular choices of $g(t)$ referred to in Propositions II and II′, yield the result that (C, ∞) *summability to* l, *whether for* s_n *or* $s(u)$, *implies* $(C, 1)$ *summability to* l *of the* (C, r) *mean*, $r \geq r_0 > 0$, *or equivalently, implies* $(C, r+1)$ *summability to* l, $r \geq r_0 > 0$. The proof of this result given here for the case of $s(u)$ appears to be the simplest possible; and a similar proof can be given for the case of s_n by first showing, adapting the arguments used for Proposition II′, that the (A) (C, r) transform or $A(x; C_r')$ tends to l as $x \to 1-0$, $r \geq r_0 > 0$, which is the incomplete form of Proposition II precisely needed for our purpose. These simple proofs, however, may be said to be indirect since they involve the (A) (C, r) transform in a role analogous to that of a catalytic agent. The alternative direct proofs not depending on the (A) (C, r) transform, given by PENNINGTON and myself ([7], [10]), are in a sense laboured and less "natural"; so that LITTLEWOOD's observation in this connexion [11a], namely, that the plausibility of a direct and natural proof "seems to fade on a closer approach", has not altogether lost its force.

In conclusion, I must record my indebtedness to Professor V. GANAPATHY IYER and Professor K. KNOPP who kindly read through my original manuscript and offered various suggestions for improvement.

References

[1] FROBENIUS, G.: Über die Leibnizsche Reihe. J. reine u. angew. Math. **89**, 262—264 (1880). — [2] HARDY, G. H.: Divergent Series. Oxford 1949. — [3] KARAMATA, J.: Über die Hardy-Littlewoodschen Umkehrungen des Abelschen Stetigkeitssatzes. Math. Z. **32**, 319—320 (1930). — [4] KARAMATA, J.: Neuer Beweis und Verallgemeinerung der Laplacesche und Stieltjessche Transformation betreffen. J. reine u. angew. Math. **164**, 27—39 (1931). — [5] KNOPP, K.: Über eine Klasse konvergenzerhaltender Integraltransformationen und den Äquivalenzsatz der C- und H-Verfahren. Math. Z. **47**, 229—264 (1941). — [6] KNOPP, K.: Eine Bemerkung zur C_k- und A-Limitierung von Funktionen. Math. Z. **50**, 155—160 (1944). — [7] PENNINGTON, W. B.: A Tauberian theorem on the oscillation of Riesz means. J. London Math. Soc. **27**, 199—206 (1952). — [8] POSTNIKOV, A. G.: A general theorem of Abelian type for power series. Doklady Akad. Nauk SSSR., N. S. **96**, 913—916 (1954) [Russian]. — [9] RAJAGOPAL, C. T.: On Riesz summability and summability by Dirichlet's series. Amer. J. Math. **76**, 252—258 (1954). — [10] RAJAGOPAL, C. T.: Note on the oscillation of Riesz, Euler and Ingham means. Quart. J. Math., Oxford II. Ser. **7**, 64—75 (1956). — [11] RAMASWAMI, V.: Some Tauberian theorems on oscillation. J. London Math. Soc. **10**, 294—308 (1935). — [11a] LITTLEWOOD, J. E.: Note on the preceding paper. J. London Math. Soc. **10**, 309—310 (1935). — [12] Sz sz, O.: Collected Mathematical Papers. Cincinnati 1955.

Ramanujan Institute of Mathematics, Madras (India)

(Eingegangen am 2. Oktober 1956)

ALBRECHT, R.
Math. Zeitschr. Bd. 67, S. 320—331 (1957)

Auswahlverfahren bei linearer Limitierung

Von
RUDOLF ALBRECHT

In einer kürzlich erschienenen Arbeit hat K. ENDL Limitierungsverfahren angegeben, die die Hausdorffschen als Spezialfall enthalten[1]). Die Verallgemeinerung besteht darin, daß aus einer zu limitierenden Folge $\{s_\mu\}$ k Teilfolgen ausgewählt werden, auf die jeweils ein Hausdorff-Verfahren angewandt wird. Eine Teilfolge enthält dabei alle Glieder s_μ, deren Indizes einer Restklasse modulo k angehören.

Da einerseits das von K. ENDL benützte Auswahlprinzip spezieller Natur ist, andererseits der Grundgedanke der Methode, die Matrix eines Limitierungsverfahrens aus Matrixelementen anderer Verfahren aufzubauen, allgemein angewandt werden kann, soll im folgenden auf solche Auswahlverfahren eingegangen werden, die aus abzählbar vielen beliebigen Teilverfahren erzeugt werden, von denen jedes auf eine bestimmte Teilfolge wirkt. Ein solches Verfahren hat den Vorteil, daß es der Struktur einer gegebenen Folge angepaßt werden kann. Sind die einzelnen Teilverfahren stets alle einer Klasse homologer Verfahren entnommen, so braucht das resultierende Verfahren nicht wieder zu dieser Klasse zu gehören. Infolgedessen lassen sich Klassen definieren, die die Klasse der Teilverfahren umfassen.

In § 1 und § 2 werden einige der praktisch wichtigsten Fragen für derartige Auswahlverfahren untersucht: Wie lassen sich die Verfahren konstruieren, wann sind sie konvergenzerhaltend, regulär, vertauschbar, verträglich, wann schließt ein Verfahren ein anderes ein? In § 3 werden als Anwendung Verfahren betrachtet, die aus Hausdorffschen gebildet sind. Schließlich soll an einem einfachen Beispiel die Wirkungsweise der Verfahren gezeigt werden.

§ 1. Konstruktion linearer Verfahren aus anderen

Die zu limitierende Folge sei $\{s_\mu\}$, dargestellt durch die einspaltige Matrix $\mathfrak{s} = (s_\mu)$. Die Menge der der Größe nach geordneten Indizes $\mu = 0, 1, 2, \ldots$ sei mit $\mathfrak{m}$ bezeichnet. Auf irgendeine Weise werde eine Zerlegung $\mathfrak{Z}(k)$ von $\mathfrak{m}$ in k, $k \leq \infty$, unendliche Teilmengen $\mathfrak{m}_\lambda$ vorgenommen, so daß

$$\mathfrak{m} = \bigcup \mathfrak{m}_\lambda, \qquad \lambda = 1, 2, \ldots, k.$$

Die Elemente von $\mathfrak{m}_\lambda$ seien mit μ_λ bezeichnet. Nach Einführung des Symbols

$$\delta_\mu^{(\lambda)} = \begin{cases} 1 & \text{falls} \quad \mu \in \mathfrak{m}_\lambda, \\ 0 & \text{andernfalls,} \end{cases}$$

[1]) ENDL, K.: Über Klassen von Limitierungsverfahren, die die Klasse der Hausdorffschen Verfahren als Spezialfall enthalten. Math. Z. 65, 113—132 (1956).

können wir

$$\mathfrak{s}_\lambda = (\delta_\mu^{(\lambda)} s_\mu)$$

setzen und haben

$$\mathfrak{s} = \sum_{\lambda=1}^{k} \mathfrak{s}_\lambda .$$

Dann ordnen wir den μ_λ eineindeutig ihre Ordnungszahlen $\varrho = \varrho(\mu_\lambda)$, $\varrho = 0, 1, 2, \ldots$, innerhalb der Menge $\mathfrak{m}_\lambda$ zu und stellen die Teilfolge $\{s_{\mu_\lambda}\}$ von $\{s_\mu\}$ durch die einspaltige Matrix

$$\mathfrak{s}_\lambda^* = (s_\varrho^{(\lambda)})$$

dar, wobei $s_{\mu_\lambda} = s_\varrho^{(\lambda)}$ gesetzt ist.

Nun soll jede der Folgen $\{s_{\mu_\lambda}\}$ einem linearen Verfahren A_λ unterworfen werden, das durch je eine Matrix $\mathfrak{A}_\lambda^* = (\alpha_{\varrho\sigma}^{(\lambda)})$ gegeben ist:

$$\mathfrak{u}_\lambda^* = \mathfrak{A}_\lambda^* \mathfrak{s}_\lambda^* \quad \text{mit} \quad \mathfrak{u}_\lambda^* = (u_\varrho^{(\lambda)}) .$$

Wir können dann aus den Matrixelementen $\alpha_{\varrho\sigma}^{(\lambda)}$ und Nullen eine Matrix $\mathfrak{A} = (a_{\mu\nu})$ aufbauen, so daß

$$\mathfrak{t} = \mathfrak{A}\mathfrak{s} \quad \text{mit} \quad \mathfrak{t} = (t_\mu)$$

und jedes t_μ gleich einem der $u_\varrho^{(\lambda)}$ ist. Wir treffen folgende Auswahl: Für $\mu = \mu_\lambda$ sei $t_{\mu_\lambda} = u_{\varrho(\mu_\lambda)}^{(\lambda)}$. Die Struktur der Matrix $\mathfrak{A}$ ist dann leicht ersichtlich.

Satz 1. Die Matrix $\mathfrak{A}$ ist die Summe von k Matrizen $\mathfrak{A}_\lambda = (a_{\mu\nu}^{(\lambda)})$, $\mathfrak{A} = \sum_{\lambda=1}^{k} \mathfrak{A}_\lambda$, wobei für das Matrixelement $a_{\mu\nu}^{(\lambda)}$ gilt

$$a_{\mu\nu}^{(\lambda)} = \begin{cases} \alpha_{\varrho(\mu_\lambda),\sigma(\nu_\lambda)}^{(\lambda)} & \text{falls } \mu = \mu_\lambda, \nu = \nu_\lambda; \mu_\lambda, \nu_\lambda \in \mathfrak{m}_\lambda; \\ 0 & \text{andernfalls.} \end{cases}$$

Beweis. Für die so gebildeten Matrizen $\mathfrak{A}_\lambda$ gilt

$$\left(\sum_{\lambda=1}^{k} \mathfrak{A}_\lambda\right) \mathfrak{s} = \sum_{\lambda=1}^{k} \mathfrak{A}_\lambda \mathfrak{s} .$$

Die einspaltige Matrix $\mathfrak{A}_\lambda \mathfrak{s}$ hat in den Zeilen $\mu \notin \mathfrak{m}_\lambda$ nur Nullen, in den Zeilen $\mu = \mu_\lambda$ stehen die Elemente

$$\sum_{\tau=0}^{\infty} a_{\mu_\lambda \tau} s_\tau = \sum_{\sigma=0}^{\infty} \alpha_{\varrho(\mu_\lambda),\sigma(\nu_\lambda)}^{(\lambda)} s_{\nu_\lambda} = u_{\varrho(\mu_\lambda)}^{(\lambda)} .$$

Das gleiche Ergebnis kommt, wenn $\mathfrak{s}$ durch $\mathfrak{s}_\lambda$ ersetzt wird: $\mathfrak{A}_\lambda \mathfrak{s}_\lambda = \mathfrak{A}_\lambda \mathfrak{s}$. Durch Addition erhalten wir

$$\mathfrak{t} = \sum_{\lambda=1}^{k} \mathfrak{A}_\lambda \mathfrak{s}_\lambda = \left(\sum_{\lambda=1}^{k} \mathfrak{A}_\lambda\right) \mathfrak{s} = \mathfrak{A}\mathfrak{s} .$$

$\mathfrak{A}$ ist die einzige Matrix der gesuchten Art, denn aus

$$\mathfrak{t} = \mathfrak{A}\mathfrak{s}, \quad \mathfrak{t} = \mathfrak{A}'\mathfrak{s}$$

folgt

$$(\mathfrak{A} - \mathfrak{A}')\mathfrak{s} = 0$$

und weil $\mathfrak{s}$ beliebig sein kann, $\mathfrak{A} = \mathfrak{A}'$.

Die Matrizen $\mathfrak{A}_\lambda$ und $\mathfrak{s}_\lambda$ sind so konstruiert, daß für sie folgende „Orthogonalitätseigenschaft" gilt:

Satz 2[2]). Für $\lambda \neq \lambda'$ ist $\tilde{\mathfrak{A}}_\lambda \mathfrak{A}_{\lambda'} = \mathfrak{A}_\lambda \tilde{\mathfrak{A}}_{\lambda'} = 0$, $\tilde{\mathfrak{s}}_\lambda \mathfrak{s}_{\lambda'} = 0$.

Beweis. Das μ, ν-Element der Produktmatrix $\mathfrak{A}_\lambda \mathfrak{A}_{\lambda'}$ ist

$$\sum_{\tau=0}^{\infty} a_{\mu\tau}^{(\lambda)} a_{\tau\nu}^{(\lambda')} = 0,$$

da $a_{\tau\nu}^{(\lambda')} = 0$ für $\tau \in \mathfrak{m}_\lambda$, $a_{\mu\tau}^{(\lambda)} = 0$ für $\tau \in \mathfrak{m}_{\lambda'}$ und $a_{\mu\tau}^{(\lambda)} = a_{\tau\nu}^{(\lambda')} = 0$ für $\tau \notin \mathfrak{m}_\lambda, \mathfrak{m}_{\lambda'}$ ist. Entsprechend gilt $\tilde{\mathfrak{A}}_\lambda \mathfrak{A}_{\lambda'} = 0$ und $\tilde{\mathfrak{s}}_\lambda \mathfrak{s}_{\lambda'} = \sum_{\tau=0}^{\infty} s_\tau^2 \delta_\tau^{(\lambda)} \delta_\tau^{(\lambda')} = 0$.

Soll umgekehrt $\mathfrak{A}_\lambda$ bzw. $\mathfrak{s}_\lambda$ aus $\mathfrak{A}$ bzw. $\mathfrak{s}$ erhalten werden, so muß $\mathfrak{A}$ bzw. $\mathfrak{s}$ mit der Matrix

$$\mathfrak{E}_\lambda = (e_{\mu\nu}^{(\lambda)}), \qquad e_{\mu\nu}^{(\lambda)} = \begin{cases} 1 & \text{falls } \mu = \nu = \mu_\lambda \\ 0 & \text{andernfalls} \end{cases}$$

multipliziert werden.

Bei gegebener Zerlegung $\mathfrak{Z}(k)$ können wir demnach das durch die Matrix $\mathfrak{A}$ bestimmte Verfahren A symbolisch als Summe der Teilverfahren A_λ schreiben:

$$A = \sum_{\lambda=1}^{k} A_\lambda.$$

Die Verfahren A_λ können aus Klassen $\mathfrak{M}_\lambda$ üblicher bekannter Verfahren entnommen sein. Es ist aber auch sinnvoll, sie in allgemeinerer Weise aus solchen zu bilden: Ist $(c_{\varkappa\sigma}^{(\lambda)})$ die Matrix eines Verfahrens aus $\mathfrak{M}_\lambda$, so kann eine Matrix $\mathfrak{A}_\lambda^* = (\alpha_{\varrho\sigma}^{(\lambda)})$ hergestellt werden, indem man *irgendeinen* Isomorphismus zwischen der Menge der Indizes ϱ und einer Teilmenge der Menge der Indizes $\varkappa$ wählt, so daß *jedem* ϱ *ein* $\varkappa$ zugeordnet ist, und man als ϱ-te Zeile von $(\alpha_{\varrho\sigma}^{(\lambda)})$ die $\varkappa$-te Zeile von $(c_{\varkappa\sigma}^{(\lambda)})$ nimmt, wenn $\varkappa$ die ϱ zugeordnete Zahl ist.

K. ENDL wählt z. B. für jedes λ für $\mathfrak{M}_\lambda$ die Klasse der Hausdorff-Verfahren, endliches k, eine Zerlegung $\mathfrak{Z}(k)$ der Indexmenge $\mathfrak{m}$ nach Restklassen modulo k und den Isomorphismus $\varrho = \varkappa$. Für $k = \infty$ und einen anderen Isomorphismus wird am Ende von § 3 ein Beispiel angegeben.

Schließlich ist zu bemerken, daß die Matrix eines Verfahrens A_λ auch aus den Matrizen mehrerer Verfahren gebildet werden kann, indem aus jeder dieser Matrizen eine Anzahl von Zeilen zur Konstruktion der Matrix von A_λ entnommen wird. Wir erläutern dies an einem einfachen Beispiel. Wir wählen $k = 1$, als zu limitierende Folge die Funktionenfolge $\{z^n\}$ und wollen diese einer regulären Transformation unterwerfen, deren Matrix aus den Matrixelementen ∞ vieler bekannter Verfahren, nämlich (E, q)-Verfahren, gebildet ist. Die Matrix eines (E, q)-Verfahrens ist bekanntlich

$$\left(\frac{\binom{\varrho}{\sigma} q^{-\sigma}}{(1+q^{-1})^\varrho} \right).$$

[2]) $\tilde{\mathfrak{A}}$ soll die transponierte Matrix von $\mathfrak{A}$ bezeichnen.

Aus jeder der ∞ vielen Matrizen für $q=0, 1, 2, \ldots$ [3]) entnehmen wir jeweils eine Zeile für die neue Matrix

$$\left(\frac{\binom{\mu^2}{\nu}\mu^{-\nu}}{(1+\mu^{-1})^{\mu^2}} \right),$$

indem wir $\varrho=\mu^2, \nu=\sigma, q=\mu$ setzen. Man erkennt leicht, daß das zugehörige Verfahren regulär ist, denn die Summe der Elemente einer Zeile der Matrix ist stets 1, der Limes der Elemente jeder Spalte ist null. Aus

$$t_\mu = \frac{\sum_{\nu=0}^{\mu^2} \binom{\mu^2}{\nu}\left(\frac{z}{\mu}\right)^\nu}{(1+\mu^{-1})^{\mu^2}} = \left(\frac{\mu+z}{\mu+1}\right)^{\mu^2}$$

ersieht man, daß die Folge $\{z^\mu\}$ in der Halbebene $\Re z<1$ vom erhaltenen Verfahren zu null limitiert wird, was keines der (E, q)-Verfahren leistet.

§ 2. Einige Sätze über Auswahlverfahren

Die bisherigen Betrachtungen waren rein formaler Art. Jetzt sollen einige Sätze angegeben werden, die Konvergenz, Regularität, Vertauschbarkeit, Verträglichkeit und Inklusionseigenschaften der zu einer Zerlegung $\mathfrak{Z}(k)$ gehörigen Verfahren $A=\sum_{\lambda=1}^{k} A_\lambda$ in Abhängigkeit vom Verhalten der Verfahren A_λ betreffen.

Wird die Folge $\{s_\mu\}$ durch A zum Wert t limitiert, dann ist klar, daß auch die Teilfolge $\{s_{\mu_\lambda}\}$ von A_λ zu t limitiert wird. Wird andererseits $\{s_{\mu_\lambda}\}$ von A_λ zu t limitiert, so kann $\{s_\mu\}$ durch A nicht zu $t'\neq t$ limitiert werden. Die Konvergenzfrage behandelt

Satz 3. Notwendig und hinreichend, damit eine Folge $\{s_\mu\}$ vom Verfahren $A=\sum_{\lambda=1}^{k} A_\lambda$ zu einem Wert t limitiert wird, ist es, daß für jeden Wert λ; $\lambda=1, 2, \ldots, k$;

(a) die Teilfolge $\{s_{\mu_\lambda}\}$ vom Teilverfahren A_λ zum Wert t limitiert wird,

(b) zu jedem Wert $\varepsilon>0$ eine von λ unabhängige Zahl M existiert, so daß $|t-t_{\mu_\lambda}|<\varepsilon$ für $\mu_\lambda>M$ ist. Ist k endlich, so entfällt (b).

Die Bedingungen sind notwendig: Gilt $t_\mu\to t$, so muß jede Teilfolge der Folge $\{t_\mu\}$ gegen t konvergieren, insbesondere jede Teilfolge $\{t_{\mu_\lambda}\}$. Wäre für Werte μ_λ Bedingung (b) nicht erfüllt, so kann auch die Konvergenzbedingung $|t-t_\mu|<\varepsilon$ für $\mu>M$ nicht gelten, da diese Werte μ_λ in der Folge der μ-Werte auftreten. Die Bedingungen sind hinreichend, da jeder der μ-Werte einem der μ_λ-Werte gleich ist und nach (b) deshalb $\{t_\mu\}$ gegen t konvergiert.

[3]) Für $q=0$ soll die Einheitsmatrix dargestellt werden.

Ist k endlich, so folgt aus der Konvergenz der Teilfolgen $|t-t_{\mu_\lambda}|<\varepsilon$ für $\mu_\lambda > M_\lambda$ die Konvergenz der Folge:

$$|t-t_\mu|<\varepsilon \quad \text{für } \mu > \operatorname{Max} M_\lambda.$$

Soll ein Verfahren A regulär sein, so genügt es nicht, daß die Teilverfahren A_λ regulär sind. Ist z.B. $k=\infty$ und hat $\mathfrak{A}_\lambda^* = (\alpha_{\varrho\sigma}^{(\lambda)})$ die Struktur

$$\alpha_{\varrho\sigma}^{(\lambda)} = \begin{cases} 0 & \text{für } \varrho \neq \sigma, \\ 1 & \text{für } \varrho = \sigma,\ \varrho > 0, \\ \lambda & \text{für } \varrho = \sigma = 0, \end{cases}$$

so sind die Verfahren A_λ gewiß regulär, die zum Verfahren $A = \sum_{\lambda=1}^{\infty} A_\lambda$ gehörige Matrix $\mathfrak{A} = (a_{\mu\nu})$ erfüllt jedoch die Toeplitzschen notwendigen und hinreichenden Bedingungen für Regularität nicht: es gibt nicht beschränkte Werte von μ, so daß

$$\sum_{\nu=0}^{\infty} |a_{\mu\nu}| = \sum_{\nu=0}^{\infty} a_{\mu\nu} = \lambda$$

ist, wobei λ mit μ gegen ∞ geht.

Satz 4. Ein Verfahren $A = \sum_{\lambda=1}^{k} A_\lambda$ ist genau dann regulär, wenn

(a) jedes Verfahren A_λ regulär ist,

(b) $\sum_{\sigma=0}^{\infty} |\alpha_{\varrho\sigma}^{(\lambda)}| < K$ und K unabhängig von ϱ *und* λ ist und $\left| \sum_{\sigma=0}^{\infty} \alpha_{\varrho\sigma}^{(\lambda)} - 1 \right| < \varepsilon$ für $\mu_\lambda > M$ mit von λ unabhängigem M gilt. Dabei bedeutet $\alpha_{\varrho\sigma}^{(\lambda)}$ das Matrixelement der zum Verfahren A_λ gehörigen Matrix $\mathfrak{A}_\lambda^*$. Ist k endlich, so entfällt (b).

Zum Beweis zeigen wir, daß für die Matrix $\mathfrak{A} = (a_{\mu\nu})$ des Verfahrens A die Toeplitzschen Bedingungen erfüllt sind: In jeder Spalte von $\mathfrak{A}$ stehen die Elemente einer Spalte einer Matrix $(\alpha_{\varrho\sigma}^{(\lambda)})$ und Nullen. Infolge der Regularität von A_λ gilt also für festes ν $a_{\mu\nu} \to 0$ für $\mu \to \infty$. In jeder Zeile von $\mathfrak{A}$ stehen die Elemente einer Zeile einer Matrix $(\alpha_{\varrho\sigma}^{(\lambda)})$ und Nullen. Infolge (b) ist für $\mu = \mu_\lambda$

$$\sum_{\nu=0}^{\infty} |a_{\mu_\lambda \nu}| = \sum_{\sigma=0}^{\infty} |\alpha_{\varrho(\mu_\lambda),\sigma}^{(\lambda)}| < K$$

unabhängig von μ_λ,

$$\sum_{\nu=0}^{\infty} a_{\mu_\lambda \nu} = \sum_{\sigma=0}^{\infty} \alpha_{\varrho(\mu_\lambda),\sigma}^{(\lambda)} \to 1 \qquad \text{für } \mu_\lambda \to \infty$$

gleichmäßig bezüglich λ.

Ist k endlich, so folgt aus der Regularität der Verfahren A_λ

$$\sum_{\sigma=0}^{\infty} |\alpha_{\varrho\sigma}^{(\lambda)}| < K_\lambda, \qquad \left| \sum_{\varrho=0}^{\infty} \alpha_{\varrho\sigma}^{(\lambda)} - 1 \right| < \varepsilon \quad \text{für} \quad \mu_\lambda > M_\lambda.$$

Dann ist (b) stets zu erfüllen, indem man $K = \operatorname{Max} K_\lambda$, $M = \operatorname{Max} M_\lambda$ setzt.

Die Notwendigkeit von (a) und (b) folgt aus der Notwendigkeit der Toeplitzschen Bedingungen.

Aus den Toeplitz-Schurschen Bedingungen erhält man weiter

Satz 5. Für $k \geq 2$ limitiert jedes konvergenzerhaltende Verfahren $A = \sum_{\lambda=1}^{k} A_\lambda$ mit der Matrix $\mathfrak{A} = (a_{\mu\nu})$ jede konvergente Folge $\{s_\mu\}$, $s_\mu \to s$, zum Wert cs mit

$$c = \lim_{\mu\to\infty} \sum_{\nu=0}^{\infty} a_{\mu\nu} = \lim_{\varrho\to\infty} \sum_{\sigma=0}^{\infty} \alpha_{\varrho\sigma}^{(\lambda)}, \qquad \lambda = 1, 2, \ldots, k.$$

Beweis. Für $k \geq 2$ enthält jede Spalte der Matrix $\mathfrak{A}$ unendlich viele Nullen. Da die Folge der Elemente jeder Spalte konvergieren muß, wenn A konvergenzerhaltend ist, muß also $\lim_{\mu\to\infty} a_{\mu\nu} = 0$ sein. Bekanntlich limitiert ein konvergenzerhaltendes Verfahren eine konvergente Folge mit dem Grenzwert s zum Wert

$$s \lim_{\mu\to\infty} \sum_{\nu=0}^{\infty} a_{\mu\nu} + \sum_{\nu=0}^{\infty} \lim_{\mu\to\infty} a_{\mu\nu}(s_\nu - s).$$

Wegen $\lim_{\mu\to\infty} a_{\mu\nu} = 0$ folgt daraus der erste Teil der Behauptung. Da $\lim_{\mu\to\infty} \sum_{\nu=0}^{\infty} a_{\mu\nu}$ existiert und für $\mu = \mu_\lambda$ $\sum_{\nu=0}^{\infty} a_{\mu_\lambda\nu} = \sum_{\sigma=0}^{\infty} \alpha_{\varrho(\mu_\lambda),\sigma}^{(\lambda)}$ ist, ist $\lim_{\mu\to\infty} \sum_{\nu=0}^{\infty} a_{\mu\nu} = \lim_{\varrho\to\infty} \sum_{\sigma=0}^{\infty} \alpha_{\varrho\sigma}^{(\lambda)}$ für jedes λ.

Satz 6. Ein Verfahren $A = \sum_{\lambda=1}^{k} A_\lambda$, $k \geq 2$, ist genau dann konvergenzerhaltend, wenn

(a) jedes Verfahren A_λ jede konvergente Folge $\{s_{\mu_\lambda}\}$, $s_{\mu_\lambda} \to s$, zum Wert cs limitiert, wobei c eine von λ unabhängige Konstante ist,

(b) $\sum_{\sigma=0}^{\infty} |\alpha_{\varrho\sigma}^{(\lambda)}| < K$ und K unabhängig von ϱ *und* λ ist und $\left| \sum_{\sigma=0}^{\infty} \alpha_{\varrho\sigma}^{(\lambda)} - c \right| < \varepsilon$ für $\mu_\lambda > M$ mit von λ unabhängigem M gilt. Bei endlichem k entfällt (b).

Nach Satz 5 gibt es für $k \geq 2$ kein konvergenzerhaltendes Verfahren A, das eine nach s konvergente Folge nicht zu einem Wert cs limitieren würde. Daraus folgt die Notwendigkeit von (a). Im übrigen ist Satz 6 wie Satz 4 zu beweisen, wobei

$$\lim_{\mu\to\infty} \sum_{\nu=0}^{\infty} a_{\mu\nu} = \lim_{\varrho\to\infty} \sum_{\sigma=0}^{\infty} \alpha_{\varrho\sigma}^{(\lambda)} = c$$

ist für jedes λ.

Satz 7. Ist $k \geq 2$ und limitiert ein Verfahren $A = \sum_{\lambda=1}^{k} A_\lambda$ jede beschränkte Folge, so kann der Limes nur null sein.

Beweis. Wie für Satz 5 folgt, daß $\lim_{\mu\to\infty} a_{\mu\nu} = 0$ ist. Da bekanntlich jedes Verfahren, das jede beschränkte Folge $\{s_\mu\}$ limitiert, diese zum Wert $\sum_{\nu=0}^{\infty} \lim_{\mu\to\infty} a_{\mu\nu} s_\nu$ limitiert, erhält man obige Behauptung.

Im folgenden seien $A=\sum_{\lambda=1}^{k} A_\lambda$, $B=\sum_{\lambda=1}^{k} B_\lambda$ zwei Verfahren, die zur selben Zerlegung $\mathfrak{Z}(k)$ gehören. Dann gelten

Satz 8. Zwei Verfahren A, B sind genau dann vertauschbar, wenn die zu gleichem λ gehörigen Teilverfahren A_λ, B_λ vertauschbar sind.

Die Bedingungen sind hinreichend: Ist $\mathfrak{A}_\lambda$ bzw. $\mathfrak{B}_\lambda$ die zur Matrix von A_λ bzw. B_λ nach § 1 gehörige Übermatrix, so daß $\mathfrak{A}=\sum_{\lambda=1}^{k}\mathfrak{A}_\lambda$, $\mathfrak{B}=\sum_{\lambda=1}^{k}\mathfrak{B}_\lambda$ ist, so muß bei Vertauschbarkeit $\mathfrak{A}\mathfrak{B}=\mathfrak{B}\mathfrak{A}$ sein. Nach Satz 2 ist

$$\mathfrak{A}\mathfrak{B}=\sum_{\lambda=1}^{k}\mathfrak{A}_\lambda\sum_{\lambda=1}^{k}\mathfrak{B}_\lambda=\sum_{\lambda=1}^{k}\mathfrak{A}_\lambda\mathfrak{B}_\lambda.$$

Da jedes Verfahren A_λ mit B_λ vertauschbar ist, gilt

$$\sum_{\lambda=1}^{k}\mathfrak{A}_\lambda\mathfrak{B}_\lambda=\sum_{\lambda=1}^{k}\mathfrak{B}_\lambda\mathfrak{A}_\lambda=\mathfrak{B}\mathfrak{A}.$$

Ist für ein λ $\mathfrak{A}_\lambda\mathfrak{B}_\lambda\neq\mathfrak{B}_\lambda\mathfrak{A}_\lambda$, so kann nicht wie vorher $\sum_{\lambda=1}^{k}\mathfrak{A}_\lambda\mathfrak{B}_\lambda=\sum_{\lambda=1}^{k}\mathfrak{B}_\lambda\mathfrak{A}_\lambda$ sein, weil die $\mathfrak{m}_\lambda$ elementenfremd sind.

Satz 9. Zwei Verfahren A, B sind verträglich, wenn für ein λ die Teilverfahren A_λ und B_λ verträglich sind.

Zwei Verfahren heißen verträglich, wenn für jede Folge $\{s_\mu\}$, für die $s_\mu\to t(A)$, $s_\mu\to t'(B)$ gilt, $t=t'$ ist. Ist also für die betrachteten Verfahren A, B $s_\mu\to t(A)$, $s_\mu\to t'(B)$ und für *ein* λ $s_{\mu_\lambda}\to t(A_\lambda)$, $s_{\mu_\lambda}\to t(B_\lambda)$, so muß $t=t'$ sein.

Satz 10. Ein Verfahren A schließt nur dann ein nicht nur zu null limitierendes Verfahren B ein, wenn für jedes λ das Verfahren A_λ das Verfahren B_λ einschließt.

Beweis. B möge eine Folge $\{s_\mu\}$ zum Wert $t\neq 0$ limitieren: $s_\mu\to t(B)$. Für $\lambda=\lambda'$ soll $A_{\lambda'}B_{\lambda'}$ *nicht* einschließen, d.h. es gibt eine Folge $\{s'_\varrho\}$, so daß $s'_\varrho\to t'(B_{\lambda'})$, aber *nicht* $s'_\varrho\to t'(A_{\lambda'})$.

Wir ersetzen die Glieder $s_{\mu_{\lambda'}}$ von $\{s_\mu\}$ durch

$$s^*_{\mu_{\lambda'}}=\frac{s'_\varrho t}{t'}\qquad\text{falls } t'\neq 0.$$

Für die so erhaltene Folge $\{s^*_\mu\}$ gilt ebenfalls $s^*_\mu\to t(B)$, aber *nicht* $s^*_\mu\to t(A)$. Ist $t'=0$, so setzen wir alle Glieder $s_\mu=0$, falls $\mu\neq\mu_{\lambda'}$, die Glieder $s_{\mu_{\lambda'}}=s'_\varrho$. Für die so konstruierte Folge $\{s^*_\mu\}$ gilt $s^*_\mu\to 0(B)$, aber *nicht* $s^*_\mu\to 0(A)$.

Korollar. Ein Verfahren A ist nur dann äquivalent einem Verfahren B, wenn für jedes λ das Verfahren A_λ dem Verfahren B_λ äquivalent ist.

Dies folgt aus der Äquivalenzdefinition und aus Satz 10.

Bei Einschränkung der Voraussetzungen läßt sich entsprechend mehr aussagen.

Satz 11. Ist k endlich, so schließt das Verfahren A genau dann das Verfahren B ein, wenn für jedes λ das Verfahren A_λ das Verfahren B_λ einschließt.

Die Bedingung ist notwendig nach Satz 10. Gilt für eine Folge $\{s_\mu\}$ $s_\mu \to t(B)$, so ist auch $s_{\mu_\lambda} \to t(B_\lambda)$ für jedes λ. A_λ schließt B_λ ein, also ist $s_{\mu_\lambda} \to t(A_\lambda)$. Aus Satz 3 folgt deshalb $s_\mu \to t(A)$.

Korollar. Bei endlichem k ist das Verfahren A genau dann äquivalent dem Verfahren B, wenn für jedes λ das Verfahren A_λ äquivalent dem Verfahren B_λ ist.

Schließlich sei noch eine Bemerkung angegeben, die sich auf die Untersuchung von Limitierungsgebieten bei Auswahlverfahren $A = \sum_{\lambda=1}^{k} A_\lambda$ bezieht. Die Glieder der zu limitierenden Folge seien nun Funktionen einer komplexen Veränderlichen z: $s_\mu = \varphi_\mu(z)$. Jedes der k Teilverfahren A_λ wirke auf eine Teilfolge $\{\varphi_{\mu_\lambda}(z)\}$ und limitiere sie in einem gewissen Gebiet G_λ der z-Ebene zu ein und derselben Funktion $\varphi(z)$, $\varphi(z) \neq \infty$, dagegen in keinem äußeren Punkt dieses Gebietes. Dann limitiert ein Verfahren $A = \sum_{\lambda=1}^{k} A_\lambda$ die Folge $\{\varphi_\mu(z)\}$ in keinem äußerem Punkt von $G = \cap G_\lambda$ zu $\varphi(z)$. Bei endlichem k wird $\{\varphi_\mu(z)\}$ in jedem inneren Punkt von G zu $\varphi(z)$ limitiert. Letzteres folgt aus Satz 3.

Sind die Gebiete G_λ die durch die Funktion $z = \psi(\zeta)$ vermittelten Bilder von (eventuell mehrblättrigen) Gebieten G_λ^* der komplexen ζ-Ebene, so gilt obige Aussage entsprechend für die Folge $\{\varphi_\mu[\psi(\zeta)]\}$ mit den Gebieten G_λ^*. Diese selbstverständliche Tatsache kann bei Auswahlverfahren nützlich verwendet werden. Besonders einfache Verhältnisse ergeben sich für die geometrische Reihe bzw. die Funktionenfolge $\{z^\mu\}$, die häufig zu Wirkfelduntersuchungen herangezogen wird. Man bestimmt dabei das Gebiet, in dem $\{z^\mu\}$ zu null limitiert wird. Wir zeigen dies an einem einfachen Beispiel:

$\{z^\mu\}$ werde in die beiden Teilfolgen $\{z^{2\sigma}\}, \{z^{2\sigma+1}\}$ zerlegt, $\sigma = 0, 1, 2, \ldots$. Auf die erste und auf die zweite wird jeweils das gleiche Verfahren, nämlich ein Borel-Verfahren mit der Matrix $(\alpha_{\varrho\sigma})$

$$\alpha_{\varrho\sigma} = \frac{e^{-\varrho}\varrho^\sigma}{\sigma!}$$

angewandt. Es ist also $A_1 = A_2$. Wir setzen $z = \zeta^{\frac{1}{2}}$. Die Folge $\{\zeta^\sigma\}$ und ebenso die Folge $\{\zeta^{\sigma+\frac{1}{2}}\}$ werden im Gebiet $\Re\zeta < 1$ zu null limitiert, für $\Re\zeta > 1$ nicht. Das resultierende Verfahren $A = A_1 + A_2$ limitiert die Folge $\{\zeta^{\mu/2}\}$ also im Gebiet G^*, der Halbebene $\Re\zeta < 1$ zu null, die Folge $\{z^\mu\}$ demnach in jedem inneren und in keinem äußeren Punkt des durch $\Re z^2 < 1$ gegebenen Gebietes G der z-Ebene. Das Verfahren A ist mit keinem Borel-Verfahren vergleichbar, denn es summiert die geometrische Reihe für Werte z mit $\Re z > 1$ und G ist im Gegensatz zu jedem Borel-Polygon nicht konvex.

Ein anderes Beispiel dieser Art ist die Summierung der geometrischen Reihe in Cassinischen Ovalen nach K. ENDL[1]), wobei die Teilverfahren Hausdorffsche sind.

§ 3. Konstruktion von Verfahren aus Hausdorff-Verfahren

Die Matrix einer Hausdorff-Transformation hat bekanntlich die Gestalt

$$\mathfrak{H} = \mathfrak{D}\,\mathfrak{M}\,\mathfrak{D}.$$

$\mathfrak{M}$ ist eine Diagonalmatrix, $\mathfrak{D}$ die Differenzenmatrix

$$\left((-1)^{\sigma}\binom{\varkappa}{\sigma}\right)$$

mit $\mathfrak{D}^2 = \mathfrak{E}$ ($\mathfrak{E}$ Einheitsmatrix). Wir wollen aus Hausdorff-Verfahren nach § 1 neue Verfahren bilden. Wir wählen eine Zerlegung $\mathfrak{Z}(k)$ der Indexmenge $\mathfrak{m}$ in k Teilmengen $\mathfrak{m}_\lambda$ ($\mu \in \mathfrak{m}$, $\mu_\lambda \in \mathfrak{m}_\lambda$), zu jeder Teilmenge $\mathfrak{m}_\lambda$ ein Hausdorff-Verfahren H_λ mit der Matrix

$$\mathfrak{D}\,\mathfrak{M}_\lambda^*\,\mathfrak{D} = (h_{\varkappa\sigma}^{(\lambda)})$$

und einen Isomorphismus

$$I_\lambda: \quad \varrho \leftrightarrow \varkappa(\varrho).$$

Damit erhalten wir Teilverfahren A_λ mit den Matrizen

$$\mathfrak{A}_\lambda^* = (\alpha_{\varrho\sigma}^{(\lambda)}) \quad \text{mit} \quad \alpha_{\varrho\sigma}^{(\lambda)} = h_{\varkappa(\varrho),\sigma}^{(\lambda)}$$

und die Matrizen

$$\mathfrak{A}_\lambda = (a_{\mu\nu}^{(\lambda)})$$

mit

$$a_{\mu\nu}^{(\lambda)} = \begin{cases} \alpha_{\varrho(\mu_\lambda),\sigma(\nu_\lambda)}^{(\lambda)} & \text{falls } \mu = \mu_\lambda,\ \nu = \nu_\lambda;\ \mu_\lambda, \nu_\lambda \in \mathfrak{m}_\lambda; \\ 0 & \text{andernfalls,} \end{cases}$$

wobei $\varrho(\mu_\lambda)$ bzw. $\sigma(\nu_\lambda)$ die Ordnungszahl von μ_λ bzw. ν_λ sei. Die Matrix des resultierenden Verfahrens A ist dann

$$\mathfrak{A} = \sum_{\lambda=1}^{k} \mathfrak{A}_\lambda.$$

Aus der Konstruktion erkennt man folgende Eigenschaften der Verfahren A_λ: Ein Verfahren A_λ ist konvergenzerhaltend bzw. regulär, wenn das Verfahren H_λ konvergenzerhaltend bzw. regulär ist. Dies folgt wieder aus der Tatsache, daß jede Teilfolge einer konvergenten Folge zum Grenzwert der Folge konvergiert.

Bekanntlich ist eine reelle Hausdorff-Transformation H_λ genau dann konvergenzerhaltend, wenn die Folge $\{m_\varkappa^{(\lambda)}\}$ der Diagonalelemente der Matrix $\mathfrak{M}_\lambda^*$ Differenz zweier total monotoner Folgen ist und genau dann regulär, wenn hierzu noch die Bedingungen

$$\lim_{\varkappa\to\infty} \Delta^\varkappa\, m_0^{(\lambda)} = 0, \qquad m_0^{(\lambda)} = 1$$

genommen werden. Im ersten Fall sind dann die $m_\varkappa^{(\lambda)}$ Momentenkonstanten vom Rang $\varkappa$

$$m_\varkappa^{(\lambda)} = \int_0^1 x^\varkappa\, d\chi^{(\lambda)},$$

wobei $\chi = \chi(x)$ von beschränkter Variation ist im Intervall $0 \leq x \leq 1$, im zweiten kommt hierzu noch die Bedingung

$$\chi(+0) = \chi(0) = 0, \quad \chi(1) = 1$$

(reguläre Momentenkonstanten). Bilden umgekehrt die $m_\varkappa^{(\lambda)}$ Momentenfolgen bzw. reguläre Momentenfolgen, so ist die zugehörige H_λ-Transformation konvergenzerhaltend bzw. regulär. In Verbindung mit den Sätzen 4 und 6 können damit Aussagen über die Regularität bzw. Konvergenzerhaltung eines Verfahrens $A = \sum_{\lambda=1}^{k} A_\lambda$ gemacht werden.

Bestehen die Isomorphismen I_λ in der Identität $\varrho = \varkappa$, so sind die Verhältnisse besonders einfach. Man kann dann

$$A_\lambda = H_\lambda,$$

$$\mathfrak{A}_\lambda^* = \mathfrak{D}\, \mathfrak{M}_\lambda^* \mathfrak{D}$$

und

$$\mathfrak{A}_\lambda = \mathfrak{D}_\lambda \mathfrak{M}_\lambda \mathfrak{D}_\lambda$$

setzen mit den Matrizen

$$\mathfrak{D}_\lambda = (d_{\mu\nu}^{(\lambda)}),\ d_{\mu\nu}^{(\lambda)} = \begin{cases} (-1)^{\sigma(\nu_\lambda)} \binom{\varrho(\mu_\lambda)}{\sigma(\nu_\lambda)} & \text{falls } \mu = \mu_\lambda, \nu = \nu_\lambda;\ \mu_\lambda, \nu_\lambda \in \mathfrak{m}_\lambda; \\ 0 & \text{andernfalls,} \end{cases}$$

$$\mathfrak{M}_\lambda = (l_{\mu\nu}^{(\lambda)}),\ l_{\mu\nu}^{(\lambda)} = \begin{cases} m_{\varrho(\mu_\lambda)}^{(\lambda)} & \text{falls } \mu = \nu = \mu_\lambda, \mu_\lambda \in \mathfrak{m}_\lambda; \\ 0 & \text{andernfalls} \end{cases}$$

und hat

$$\mathfrak{A} = \sum_{\lambda=1}^{k} \mathfrak{D}_\lambda \mathfrak{M}_\lambda \mathfrak{D}_\lambda.$$

Wie in Satz 2 kann man zeigen, daß für $\lambda \neq \lambda'$

$$\mathfrak{D}_\lambda \mathfrak{D}_{\lambda'} = 0$$

und

$$\mathfrak{D}_\lambda \mathfrak{M}_{\lambda'} = \mathfrak{M}_{\lambda'} \mathfrak{D}_\lambda = 0$$

ist und weiter infolge $\mathfrak{D}^2 = \mathfrak{E}$

$$\mathfrak{D}_\lambda^2 = \mathfrak{E}_\lambda \quad \text{mit} \quad \sum_{\lambda=1}^{k} \mathfrak{E}_\lambda = \mathfrak{E}$$

gilt. Daraus folgt

Satz 12. Eine Transformation $A = \sum_{\lambda=1}^{k} H_\lambda$ hat die Matrix $\mathfrak{A} = \mathfrak{D}_0 \mathfrak{M} \mathfrak{D}_0$ mit $\mathfrak{D}_0 = \sum_{\lambda=1}^{k} \mathfrak{D}_\lambda$, $\mathfrak{M} = \sum_{\lambda=1}^{k} \mathfrak{M}_\lambda$, wobei $\mathfrak{D}_0^2 = \mathfrak{E}$ und $\mathfrak{M}$ eine Diagonalmatrix ist.

Für endliches k können für die Transformationen $A = \sum_{\lambda=1}^{k} H_\lambda$ mit Hilfe der Ergebnisse von § 2 und der oben erwähnten Eigenschaften von reellen

Hausdorff-Transformationen eine Reihe von Sätzen angegeben werden. Zum Beispiel gilt

Satz 13. Für endliches k, $k \geq 2$, ist ein Verfahren $A = \sum_{\lambda=1}^{k} H_\lambda$ genau dann konvergenzerhaltend bzw. regulär, wenn die zu den Hausdorff-Verfahren H_λ gehörigen Konstanten $m_\varrho^{(\lambda)}$ eine Momentenfolge bilden, für jedes λ $\lim_{\varrho\to\infty} \Delta^\varrho m_0^{(\lambda)} = 0$ und $m_0^{(\lambda)} = c$, c eine Konstante, bzw. $c = 1$ ist.

Nach Satz 6 ist bei endlichem k, $k \geq 2$, A genau dann konvergenzerhaltend, wenn jedes Teilverfahren H_λ jede konvergente Folge $s_{\mu_\lambda} \to s$ zum Wert cs limitiert. Bekanntlich ist ein Hausdorff-Verfahren H_λ genau dann konvergenzerhaltend, wenn die $m_\varrho^{(\lambda)}$ eine Momentenfolge bilden. Ferner limitiert es zum Wert

$$m_0^{(\lambda)} s + \lim_{\varrho\to\infty} \Delta^\varrho m_0^{(\lambda)} (s_0 - s).$$

Für ein konvergenzerhaltendes bzw. reguläres Verfahren A ist deshalb

$$\lim_{\varrho\to\infty} \Delta^\varrho m_0^{(\lambda)} = 0, \qquad m_0^{(\lambda)} = c \text{ bzw. } 1$$

notwendig und hinreichend.

Weiter kann man leicht aus der Vertauschbarkeit und Verträglichkeit von Hausdorff-Verfahren H_λ auf die entsprechenden Eigenschaften für diejenigen der von uns betrachteten Verfahren schließen, die zur gleichen Zerlegung $\mathfrak{Z}(k)$ gehören. Diese und weitere hierher gehörige Sätze sollen jedoch nicht mehr formuliert werden, da sie sich für spezielle Zerlegungen $\mathfrak{Z}(k)$ bereits bei K. ENDL[1]) finden. Auch wird dort ein interessantes Beispiel für die Wirkungsweise dieser Verfahren mit endlichem k gegeben. Wir wollen statt dessen wieder zum allgemeinen Fall mit beliebigen Isomorphismen I_λ zurückkehren und noch ein einfaches Beispiel für $k = \infty$ betrachten:

Gegeben ist die Folge

$$2^0\ 0\ 0\ 1\ 1\ 2^1\ 0\ 0\ 0\ 0\ 1\ 1\ 1\ 1\ 2^2\ 0\ \ldots.$$

Auf $(2\tau+1)$ von null verschiedene Zahlen, von denen die $(2\tau+1)$-te 2^τ, die übrigen 1 sind, folgen $2(\tau+1)$ Nullen; $\tau = 0, 1, 2, \ldots$. Wir wollen die Folge durch ein reguläres Verfahren $A = \sum_{\lambda=1}^{\infty} A_\lambda$ limitieren.

Wir führen eine Zerlegung $\mathfrak{Z}(\infty)$ in Teilfolgen durch, indem wir aus obiger Folge die erste von null verschiedene Zahl, die erste dieser Zahl nachfolgende Null, dann wieder die erste nachfolgende von null verschiedene Zahl usw. herausgreifen, aus der verbleibenden Folge die nächste Teilfolge entsprechend bilden, jedoch mit der ersten Null beginnen, und so fortfahren. Für die λ-te Teilfolge erhält man

$$0\ 1\ 0\ 1\ 0\ 1\ \ldots \quad \text{falls } \lambda = 2\tau,$$
$$2^\tau 0\ 1\ 0\ 1\ 0\ \ldots \quad \text{falls } \lambda = 2\tau + 1$$

ist. Die Matrix $\mathfrak{A}_\lambda^* = (\alpha_{\varrho\sigma}^{(\lambda)})$ des Verfahrens A_λ, das auf die λ-te Teilfolge wirkt, bilden wir aus der Dreiecksmatrix $(h_{\varkappa\sigma})$

$$h_{\varkappa\sigma} = \begin{cases} \frac{1}{1+\varkappa} & \text{falls } \sigma \leq \varkappa, \\ 0 & \text{andernfalls} \end{cases}$$

eines $(C, 1)$-Verfahrens. Die Isomorphismen I_λ: $\varrho \leftrightarrow \varkappa$ wählen wir so, daß

$$\begin{aligned} \varkappa &= \mu_\lambda \quad \text{für} \quad \varrho = \varrho(\mu_\lambda), \quad \text{falls} \quad \lambda = 2\tau, \\ \varkappa &= \mu_\lambda 2^\tau \quad \text{für} \quad \varrho = \varrho(\mu_\lambda), \quad \text{falls} \quad \lambda = 2\tau + 1 \end{aligned}$$

ist. $\varrho(\mu_\lambda)$ soll die Ordnungszahl von μ_λ bezeichnen. Aus der Regularität des $(C, 1)$-Verfahrens folgt die Regularität der Verfahren A_λ. Die Summe der Elemente einer Zeile ist stets 1. Also ist das resultierende Verfahren $A = \sum_{\lambda=1}^{\infty} A_\lambda$ nach Satz 4 regulär. Ferner wird jede der angegebenen Teilfolgen $(C, 1)$-limitiert zum Wert $\frac{1}{2}$. Durch die Wahl der Isomorphismen I_λ wird erreicht, daß

$$\frac{1}{2} - t_\mu = O\left(\frac{1}{\mu}\right),$$

also wird die Gesamtfolge vom Verfahren A zu $\frac{1}{2}$ limitiert. Das Verfahren A ist aus $(C, 1)$-Verfahren gebildet, die angegebene Folge ist jedoch durch kein (C, k)-Verfahren limitierbar, weil in diesem Falle

$$s_\mu = o(\mu^k)$$

sein müßte, während in unserer Folge Glieder

$$s_{2\tau^2+3\tau} = 2^\tau$$

auftreten.

Zusatz bei der Korrektur. Es sei noch bemerkt, daß die von HAUSDORFF gegebene Verallgemeinerung der nach ihm benannten Verfahren [HAUSDORFF, F.: Summationsmethoden und Momentenfolgen. II. Math. Z. **9**, 280—299 (1921)] von der hier betrachteten verschieden ist. Sie besteht im wesentlichen darin, daß für die Ordnung der Momentenkonstanten die Zahlen einer beliebigen (geeigneten) Folge zugelassen werden, während hier stets Momentenkonstanten der Ordnung $\varkappa = 0, 1, 2, \ldots$ auftreten.

Mathematisches Institut der Technischen Hochschule München

(Eingegangen am 10. Oktober 1956)

LINGENBERG, R.
Math. Zeitschr. Bd. 67, S. 332—360 (1957)

Zur Einführung von Koordinaten in einer projektiven Ebene mit Hilfe von Endomorphismen transitiver Translationsgruppen

Von

ROLF LINGENBERG

In der Theorie der projektiven Ebenen steht an zentraler Stelle das Problem, wie man eine durch die üblichen drei projektiven Inzidenzaxiome gegebene projektive Ebene Π als Koordinatenebene über einem algebraischen System darstellen kann, um dann bei speziellen Fragen die Vorteile des algebraischen Rechnens im Koordinatenbereich ausnutzen zu können. Zur Lösung dieser Aufgabe sind verschiedene Wege eingeschlagen worden, so von HILBERT [9], SCHWAN [14], ARTIN [3], BAER [4] und HALL [6][1]), um nur einige zu nennen. Die Koordinateneinführung nach ARTIN, die inhaltlich im wesentlichen auf SCHWAN [14] zurückgeht, steht hierbei den heute üblichen Begriffsbildungen und Methoden am nächsten und ist wohl auch das eleganteste Verfahren. Der Koordinatenbereich wird dabei als Unterring $\mathfrak{E}$ des Endomorphismenringes der abelschen Translationsgruppe der affinen Ebene gewonnen, die man aus Π durch Auszeichnen einer Gerade u als uneigentlicher Geraden erhält. $\mathfrak{E}$ besteht dann aus Automorphismen der Translationsgruppe, die durch Transformieren der Translationen mit den Streckungen der affinen Ebene entstehen, und dem Endomorphismus 0, der alle Translationen in das Einselement abbildet. $\mathfrak{E}$ ist ein Schiefkörper. Die Möglichkeit jedem Punkt der affinen Ebene genau ein geordnetes Paar (ξ, η) von Endomorphismen $\in \mathfrak{E}$ zuzuordnen, beruht dann auf der linearen Transitivität jeder Gruppe der genannten Streckungen. Der Nachteil dieser Koordinateneinführung besteht darin, daß sie nur für eine spezielle Klasse von projektiven Ebenen durchführbar ist, nämlich für diejenigen projektiven Ebenen, in denen es eine transitive Translationsgruppe und eine linear transitive Gruppe von Streckungen, also von Homologien in Π mit der gleichen Achse wie die Translationen gibt.

In der vorliegenden Arbeit wird nun versucht, eine analoge Koordinateneinführung auch für die projektiven Ebenen zu entwickeln, die die erwähnten Transitivitätseigenschaften nicht besitzen. Dabei werden die Translationen und Homologien durch Abbildungen einer Geraden x (der „*x-Achse*") auf sich ersetzt, die sich als Produkt zweier Perspektivitäten (Definition s. § 1) schreiben lassen und die sich in die Translationen bzw. Homologien einbetten lassen — sofern diese existieren sollten (§ 3). Man gewinnt dadurch Mengen $\mathfrak{T}$ und $\mathfrak{H}$

[1]) Zahlen in eckigen Klammern verweisen auf das Literaturverzeichnis.

von Abbildungen, die genau einfach transitiv auf der Menge der Punkte von x, also linear transitiv sind (abgesehen von Ausnahmepunkten) und mit deren Hilfe man die Artinsche Endomorphismenrechnung nachahmen kann, wenn man statt der Menge $\mathfrak{E}$ eine Menge $\mathfrak{A}$ von Abbildungen der Menge $\mathfrak{T}$ auf sich betrachtet, die durch Verwendung der Elemente $\in \mathfrak{H}$ erklärt werden (§ 4). Im speziellen — also bei Voraussetzung weiterer Axiome für Π — hat $\mathfrak{A}$ entsprechende Eigenschaften wie $\mathfrak{E}$: Ist z.B. die Translationsgruppe transitiv, so kann man $\mathfrak{T}$ durch die Gruppe aller Translationen in *„x-Richtung"* ersetzen und zeigen, daß die Elemente $\in \mathfrak{A}$ dann Endomorphismen dieser Gruppe sind (§ 5).

Das durch das beschriebene Verfahren erhaltene System $\mathfrak{A}$, das dann als Koordinatenbereich für die gegebene Ebene Π verwendet wird, ist natürlich im allgemeinen kein Schiefkörper. Assoziativ-, Kommutativ- und Distributivgesetze brauchen nicht zu gelten. Die Gültigkeit jedes dieser Gesetze hängt an der Voraussetzung weiterer Axiome für Π, z.B. ist die Tatsache, daß die Elemente $\in \mathfrak{A}$ Endomorphismen sind, gleichwertig mit der Gültigkeit des Linksdistributivgesetzes, so daß dieses erst dann erschlossen werden kann, wenn die volle Translationsgruppe transitiv ist. Es ergeben sich dann sehr einfach eine ganze Reihe von Äquivalenzen zwischen den projektiven Ebenen, in denen geeignete Zusatzaxiome gelten, und den projektiven Koordinatenebenen über algebraischen Systemen, in denen das eine oder andere der erwähnten Rechengesetze zusätzlich gilt. Von diesen Äquivalenzen, die den bisher bei anderen Koordinateneinführungen gewonnenen gleichwertig sind[2]), werden einige wichtige Stufen dargestellt, und zwar die Kennzeichnung der projektiven Ebenen über cartesischen Zahlensystemen, Veblen-Wedderburn-Systemen, Fastkörpern, Schiefkörpern und kommutativen Körpern (§ 5)[3]).

Die zur Kennzeichnung der ersten vier Stufen verwandten Zusatzaxiome enthalten — gemäß der gewählten Methode — ausschließlich Forderungen über die Transitivität von Gruppen perspektiver Kollineationen (also von Translations- und Homologiegruppen)[4]). Die letzte Stufe wird, einer Arbeit von BAER [*4*] folgend, zunächst durch die Existenz einer perspektiven Dualität gekennzeichnet. Es wird dann aber darüber hinaus noch gezeigt, daß man statt dessen auch die schwächere Forderung der Existenz einer quasiprojektiven Dualität stellen kann. Um die Beziehungen zwischen den perspektiven Abbildungen in Π (also den Perspektivitäten, perspektiven Kollineationen und Dualitäten) deutlich werden zu lassen, die den Äquivalenzen von Rechengesetz und den benutzten Zusatzforderungen an Π zugrunde liegen, wurde

[2]) Vgl. dazu das Buch von G. PICKERT [*13*].

[3]) Hierbei ist zu sagen, daß jede dieser Stufen nicht mit der nächstfolgenden zusammenfällt, da man für jede Stufe echte algebraische Systeme kennt, die zwar die der Stufe entsprechenden Rechengesetze, aber nicht die der nächstfolgenden Stufe erfüllen.

[4]) Bei den übrigen Arten von Koordinateneinführungen werden als Zusatzaxiome meistens die Gültigkeit von Schließungssätzen gefordert, so etwa von Spezialfällen des Satzes von DESARGUES, die im wesentlichen der Transitivität von speziellen Gruppen perspektiver Kollineationen äquivalent ist. Vgl. [*13*].

darauf verzichtet, mit Koordinaten zu rechnen*). Außerdem stellten sich bei diesem Vorgehen für sich interessante Tatsachen über die perspektiven Abbildungen heraus (vgl. dazu auch Satz 16 und 17 aus dem Anhang, in dem ein in § 5 geführter Beweis bei schwächeren Voraussetzungen wiederholt wird).

Für den Beweis, daß aus den Eigenschaften der Automorphismen $\in \mathfrak{A}$, die den Rechengesetzen gleichwertig sind, die Existenz von perspektiven Kollineationen und Dualitäten folgt, ist ein allgemeiner Satz von besonderer Bedeutung, der in § 2 bewiesen wird und im wesentlichen besagt, daß ein Isomorphismus zwischen einfach transitiven Gruppen perspektiver Kollineationen, der die Gruppen komponentenweise ineinander überführt, stets durch einen inneren Automorphismus der Gruppe aller Kollineationen und Dualitäten der projektiven Ebene induziert wird.

§ 1. Perspektive Kollineationen und perspektive Dualitäten

Wir verstehen unter einer *projektiven Ebene* Π eine Menge von Punkten und Geraden, in der eine Inzidenz zwischen Punkt und Gerade erklärt ist und in der folgende Axiome gelten:

P1. *Zu zwei verschiedenen Punkten P, Q gibt es eine und nur eine Gerade $g=(P, Q)$, die mit P und Q inzidiert (geschrieben $g \mathrel{I} P, Q$).*

P2. *Zu zwei verschiedenen Geraden g, h gibt es einen und nur einen Punkt $P=(g, h)=g \cdot h$, der mit g und h inzidiert.*

P3. *Es gibt wenigstens vier Punkte, von denen keine drei mit einer Geraden inzidieren.*

Aus Axiom P3 folgt, daß auf jeder Geraden wenigstens drei verschiedene Punkte liegen. Um Ausnahmefälle zu vermeiden, wollen wir darüber hinaus — soweit notwendig — voraussetzen, daß auf jeder Geraden wenigstens vier verschiedene Punkte liegen (z.B. für Satz 1—3, Lemma 12, 13).

Unter einer *Perspektivität* in der projektiven Ebene verstehen wir eine eineindeutige Abbildung der Punkte auf einer Geraden g auf die Menge der Punkte einer Geraden h, bei der Punkt und Bildpunkt jeweils auf einer Geraden durch einen festen Punkt Z liegen, der nicht mit g oder h inzidiert[5]). Wir bezeichnen eine Perspektivität mit π_{gZh}. Der Punkt Z heißt das Zentrum von π_{gZh}.

Eine eineindeutige Abbildung der Menge der Punkte einer projektiven Ebene auf sich heißt eine *Kollineation*, wenn kollineare Punkte (d.h. Punkte auf einer Geraden) in kollineare Punkte übergehen; und eine eineindeutige Abbildung der Menge der Punkte der projektiven Ebene auf die Menge der Geraden der projektiven Ebene heißt eine *Dualität*[6]), wenn kollineare Punkte in konkurrente Geraden (d.h. Geraden durch einen Punkt) übergehen.

---

*) wodurch die Beweise auch nur zum Teil abgekürzt worden wären.

[5]) Geschrieben $Z \nmid g, h$.

[6]) Auch Korrelation genannt.

Eine Kollineation π heißt eine *perspektive Kollineation*[7]) *mit Zentrum Z und Achse a*, wenn

$$P\pi = P \quad \textit{für alle } P \mathrel{I} a, \qquad g\pi = g \quad \textit{für alle } g \mathrel{I} Z$$

gilt. Ist $Z \mathrel{I} a$, so heißt π eine *Translation*[8]) und ist $Z \mathrel{\not I} a$, so heißt π eine *Homologie*.

Eine Dualität δ einer projektiven Ebene heißt eine *perspektive Dualität mit Zentrum Z und Achse a*, wenn für $Z \mathrel{\not I} a$

$$P\delta = (P, Z) \quad \textit{für alle } P \mathrel{I} a, \qquad g\delta = (g, a) \quad \textit{für alle } g \mathrel{I} Z$$

gilt.

Ist π eine perspektive Kollineation mit Z als Zentrum und a als Achse, so gilt für alle Punkte P und alle Geraden g:

(1) *$P, P\pi, Z$ sind kollinear und $g, g\pi, a$ sind konkurrent.*

Für die perspektiven Kollineationen und Dualitäten gilt bekanntlich weiter[9]):

Lemma 1. *Es gibt höchstens eine perspektive Kollineation mit Z als Zentrum und a als Achse, welche einen Punkt P auf einen Punkt Q auf (P, Z) $[P, Q \neq Z;$ $P, Q \mathrel{\not I} a]$ abbildet, und höchstens eine perspektive Dualität mit Z als Zentrum und a als Achse, welche einen Punkt P $[P \mathrel{\not I} a;$ $P \neq Z]$ auf eine Gerade q $[q \neq a,$ $(P, Z)]$ durch $(P, Z) \cdot a$ abbildet.*

Lemma 2. *Besitzt eine Kollineation α entweder eine Fixgerade a, die punktweise festbleibt, oder einen Fixpunkt Z, der geradenweise festbleibt, so ist α eine perspektive Kollineation mit a als Achse bzw. mit Z als Zentrum.*

Beweis von Lemma 2. Da die zweite Aussage von Lemma 2 dual zu der ersten ist, genügt es, diese zu beweisen. Es möge also eine Gerade a geben, die bei α punktweise festbleibt. Dann gilt:

(2) *Ist $P \neq P\alpha$, so ist $(P, P\alpha) = g$ Fixgerade von α.*

Denn sei $(g, a) = Q$, also $g = (P, Q) = (P\alpha, Q)$, so gilt $Q\alpha = Q$ und damit $g\alpha = (P, Q)\alpha = (P\alpha, Q) = g$. Da nun der Schnittpunkt zweier Fixgeraden Fixpunkt ist und da ferner eine Kollineation $\neq 1$, die eine Gerade a punktweise festläßt, höchstens einen Fixpunkt Z mit $Z \mathrel{\not I} a$ besitzen kann, folgt aus (2):

(3) *Es gibt einen Punkt Z, so daß für alle P mit $P \neq P\alpha$ gilt $Z \mathrel{I} (P, P\alpha)$. Z ist Fixpunkt von α.*

Da nun jede Gerade g durch Z Fixgerade ist [wählt man nämlich einen Punkt P auf g mit $P \neq P\alpha$, so gilt nach (3) $(P, P\alpha) \mathrel{I} Z$ und damit $(P, P\alpha) = g$], muß α perspektive Kollineation mit a als Achse und Z als Zentrum sein.

[7]) In [13] zentrale Kollineation genannt.

[8]) Dieser Name wurde gewählt, weil die Translationen in einer affinen Ebene, als Kollineationen der projektiven Erweiterung derselben aufgefaßt, perspektive Kollineationen sind, deren Zentrum auf der Achse (nämlich der uneigentlichen Geraden) liegt.

[9]) Siehe [4] oder [13].

Ist α eine beliebige Kollineation oder Dualität und ist π eine perspektive Kollineation bzw. Dualität mit Z als Zentrum und a als Achse, so ist offenbar $\alpha^{-1}\pi\alpha$ perspektive Kollineation bzw. Dualität mit $Z\alpha$ als Zentrum und $a\alpha$ als Achse, wenn α Kollineation, und mit $a\alpha$ als Zentrum und $Z\alpha$ als Achse, wenn α Dualität ist.

§ 2. Transitive Gruppen von perspektiven Kollineationen

Die perspektiven Kollineationen mit festem Zentrum Z und fester Achse a bilden eine Gruppe, die wir mit $\mathfrak{P}_{Za}$ bezeichnen. Die Menge $\mathfrak{D}_{Za}$ der perspektiven Dualitäten mit festem Zentrum Z und fester Achse a erzeugt (sofern sie nicht leer ist) eine Gruppe, die die Gruppe $\mathfrak{P}_{Za}$ der Homologien mit Z als Zentrum und a als Achse als Untergruppe vom Index 2 und $\mathfrak{D}_{Za}$ als deren Nebenklasse enthält.

Die Existenz von perspektiven Kollineationen und Dualitäten ist für eine beliebige projektive Ebene selbstverständlich nicht gesichert; die Gruppen $\mathfrak{P}_{Za}$ können also alle nur aus dem Einselement bestehen und die Mengen $\mathfrak{D}_{Za}$ alle leer sein. Wir definieren nun:

Eine Gruppe $\mathfrak{P}_{Za}$ von perspektiven Kollineationen heißt *linear transitiv* auf einer festen Geraden g mit $g \mathrel{I} Z$, wenn es zu zwei beliebigen Punkten P, Q auf g $[P, Q \neq Z;\ P, Q \nmid a]$ eine perspektive Kollineation $\pi \in \mathfrak{P}_{Za}$ mit $P\pi = Q$ gibt. Entsprechend heißt die Menge $\mathfrak{D}_{Za}$ *linear transitiv* auf einer festen Geraden g mit $g \mathrel{I} Z$, wenn es zu einem beliebigen Punkt P auf g $[P \neq Z;\ P \nmid a]$ und einer beliebigen Geraden q durch $g \cdot a$ $[q \neq a, g]$ eine perspektive Dualität $\delta \in \mathfrak{D}_{Za}$ mit $P\delta = q$ gibt.

Man sieht leicht ein[10]), daß $\mathfrak{P}_{Za}$ bzw. $\mathfrak{D}_{Za}$ dann und nur dann auf einer Geraden g linear transitiv ist, wenn sie auch auf einer beliebigen anderen Geraden g' mit $g' \mathrel{I} Z$ linear transitiv ist. Den Zusatz *„auf einer Geraden g"* zu linear transitiv können wir daher hinfort weglassen. Ebenfalls ohne Mühe zeigt man[10]) das folgende

Lemma 3. *Ist die Gruppe $\mathfrak{P}_{Za}$ und auch die Gruppe $\mathfrak{P}_{Z'a}$ mit $Z' \neq Z$ linear transitiv, so ist jede Gruppe $\mathfrak{P}_{Z''a}$ mit $Z'' \mathrel{I} (Z, Z')$ linear transitiv.*

Die Voraussetzungen von Lemma 3 können im allgemeinen noch abgeschwächt werden:

Lemma 4. *Ist die Gruppe $\mathfrak{P}_{Za}$ linear transitiv und gibt es eine Kollineation α mit $a\alpha = a$, jedoch $Z\alpha \neq Z$, so ist jede Gruppe $\mathfrak{P}_{Z'a}$ mit $Z' \mathrel{I} (Z, Z\alpha)$ linear transitiv.*

Der Beweis von Lemma 4 ergibt sich sofort aus Lemma 3, wenn man bedenkt, daß mit $\mathfrak{P}_{Za}$ auch $\alpha^{-1}\mathfrak{P}_{Za}\alpha$ eine linear transitive Gruppe von perspektiven Kollineationen ist, deren Elemente a als Achse und $Z\alpha \neq Z$ als Zentrum haben.

Die Menge aller perspektiven Kollineationen mit gleicher Achse u bildet nach Lemma 2 eine Gruppe $\mathfrak{P}_u$. Sind $\pi, \pi' \neq 1$ zwei Elemente $\in \mathfrak{P}_u$ und Z, Z' deren Zentren, so liegt das Zentrum der perspektiven Kollineation $\pi\pi'$

[10]) Siehe BAER [4].

entweder auf (Z, Z') — für $Z \neq Z'$ — oder es ist gleich Z, wenn nämlich $Z = Z'$ ist. Daher bilden alle diejenigen Elemente $\in \mathfrak{P}_u$, deren Zentren auf einer festen Geraden v liegen, eine Untergruppe von $\mathfrak{P}_u$, die wir mit $\mathfrak{P}_{vu}$ bezeichnen wollen.

Dual hierzu bilden alle perspektiven Kollineationen mit gleichem Zentrum U eine Gruppe $\mathfrak{P}_U$, die als Untergruppe $\mathfrak{P}_{UV}$ die Menge aller Elemente $\in \mathfrak{P}_U$ mit einer Achse durch einen festen Punkt V enthält. $\mathfrak{P}_{uu}$ ist ersichtlich die Gruppe aller Translationen mit u als Achse und $\mathfrak{P}_{UU}$ die Gruppe aller Translationen mit U als Zentrum.

Wir nennen nun jede Untergruppe $\mathfrak{P}_{Zu}$ von $\mathfrak{P}_{vu}$ [für $Z \mathrel{I} v$] und entsprechend jede Untergruppe $\mathfrak{P}_{Ua}$ von $\mathfrak{P}_{UV}$ [für $a \mathrel{I} V$] eine *Komponente* von $\mathfrak{P}_{vu}$ bzw. von $\mathfrak{P}_{UV}$. Sind zwei Komponenten der Gruppe $\mathfrak{P}_{vu}$ bzw. der Gruppe $\mathfrak{P}_{UV}$ linear transitiv, so ist nach Lemma 3 (bzw. der dualen Aussage) jede Komponente von $\mathfrak{P}_{vu}$ bzw. $\mathfrak{P}_{UV}$ linear transitiv. Die Gruppe $\mathfrak{P}_{vu}$ (und entsprechend die Gruppe $\mathfrak{P}_{UV}$) ist dann als Produkt (im Sinne der Komplexmultiplikation in einer Gruppe) zweier Komponenten darstellbar. Denn sei Z ein Punkt mit $Z \mathrel{I} v$, u und Z' ein Punkt $\neq Z$ auf v, ferner π ein beliebiges Element $\neq 1$ mit $\pi \in \mathfrak{P}_{vu}$ und endlich P ein beliebiger Punkt mit $P \nmid v, u$, so ist $(P, Z') \neq (P\pi, Z)$ und $Q = (P, Z') \cdot (P\pi, Z) \nmid v, u$, außerdem aber natürlich auch $P\pi \nmid v, u$. Mithin gibt es ein $\omega \in \mathfrak{P}_{Z'u}$ mit $P\omega = Q$ und ein $\tau \in \mathfrak{P}_{Zu}$ mit $Q\tau = P\pi$. Damit gilt $P\omega\tau = P\pi$ und, da $\omega\tau$ und π dasselbe Zentrum haben müssen (nämlich $Z'' = (P\pi, P) \cdot v$), nach Lemma 1 weiter $\pi = \omega\tau$. Also folgt $\mathfrak{P}_{vu} = \mathfrak{P}_{Z'u}\mathfrak{P}_{Zu}$. Entsprechendes erhält man für die Gruppe $\mathfrak{P}_{UV}$ mit dualen Schlüssen.

Die Darstellung $\pi = \omega\tau$ von $\pi \in \mathfrak{P}_{vu}$ als Produkt einer perspektiven Kollineation $\omega \in \mathfrak{P}_{Z'u}$ und einer Translation $\tau \in \mathfrak{P}_{Zu}$ ist eindeutig, denn wäre $\pi = \omega'\tau'$ mit $\omega' \in \mathfrak{P}_{Z'u}$ und $\tau' \in \mathfrak{P}_{Zu}$, so folgte aus $\omega \neq \omega'$ auch $P\omega \neq P\omega'$, also $(P\omega, Z) \neq (P\omega', Z)$. Dann aber läge $P\omega'\tau' = P\pi$ auf $(P\omega', Z)$ und auf $(P\omega, Z)$, wäre also gleich Z, im Widerspruch zu $P\pi \nmid v, u$. Also gilt $\omega = \omega'$ und damit wegen $P\omega\tau' = P\pi = P\omega\tau$ nach Lemma 1 $\tau' = \tau$. Dual hierzu muß gelten, daß sich jedes Element der Gruppe $\mathfrak{P}_{UV}$ eindeutig als Produkt $\omega\tau$ mit $\omega \in \mathfrak{P}_{Ua}$ und $\tau \in \mathfrak{P}_{Ua'}$ $[a, a' \mathrel{I} V;\ a' \mathrel{I} U;\ a \neq a']$ darstellen läßt.

Die Gruppe $\mathfrak{P}_{vu}$ ist dann genau einfach transitiv[11]) auf der Menge der Punkte P mit $P \nmid u, v$, und die Gruppe $\mathfrak{P}_{UV}$ ist dann genau einfach transitiv auf der Menge der Geraden g mit $g \nmid U, V$. Wir bezeichnen das eindeutig bestimmte Element $\in \mathfrak{P}_{vu}$, welches den Punkt P in den Punkt Q überführt $[P, Q \nmid u, v]$, mit π_{PQ}; entsprechend mit π_{pq} das eindeutig bestimmte Element $\in \mathfrak{P}_{UV}$, welches die Gerade p in die Gerade q überführt $[p, q \nmid U, V]$. Für die Transitivität der Gruppe $\mathfrak{P}_{vu}$ bzw. der Gruppe $\mathfrak{P}_{UV}$ genügt nach Lemma 4 im Falle $u \neq v$ bzw. $U \neq V$ die Voraussetzung, daß eine einzige Komponente $\mathfrak{P}_{Zu}$ mit $Z \nmid u$ linear transitiv ist und eine andere nicht nur aus dem Einselement besteht.

---

[11]) Eine Menge $\mathfrak{M}$ von Abbildungen in einer projektiven Ebene heißt *einfach transitiv* auf einer Menge $\mathfrak{N}$ von Punkten bzw. Geraden, wenn es zu zwei Elementen $A, B \in \mathfrak{N}$ stets eine Abbildung $\alpha \in \mathfrak{M}$ mit $A\alpha = B$ gibt.

Ist α eine Kollineation, so ist die Abbildung φ mit $\pi^\varphi = \alpha^{-1}\pi\alpha$ für alle $\pi \in \mathfrak{P}_{vu}$ ein Isomorphismus der Gruppe $\mathfrak{P}_{vu}$ auf die Gruppe $\mathfrak{P}_{v'u'}$ mit $u' = u\alpha$ und $v' = v\alpha$, der eine Komponente von $\mathfrak{P}_{vu}$ in eine Komponente von $\mathfrak{P}_{v'u'}$ überführt. Umgekehrt gilt

Satz 1. *Sind die Gruppen $\mathfrak{P}_{vu}$ und $\mathfrak{P}_{v'u'}$ einfach transitiv*[12]) *und ist φ ein Isomorphismus von $\mathfrak{P}_{vu}$ und $\mathfrak{P}_{v'u'}$, der jede Komponente von $\mathfrak{P}_{vu}$ in eine Komponente von $\mathfrak{P}_{v'u'}$ überführt, so gibt es zu jedem Punktepaar P, P' $[P \nmid u, v;$ $P' \nmid u', v']$ eine Kollineation α, so daß*

$$\pi^\varphi = \alpha^{-1}\pi\alpha \qquad \text{für alle } \pi \in \mathfrak{P}_{vu}$$

und $P\alpha = P'$ gilt.

Satz 2. *Sind die Gruppen $\mathfrak{P}_{UV}$ und $\mathfrak{P}_{U'V'}$ einfach transitiv*[12]) *und ist φ ein Isomorphismus von $\mathfrak{P}_{UV}$ und $\mathfrak{P}_{U'V'}$, der jede Komponente von $\mathfrak{P}_{UV}$ in eine Komponente von $\mathfrak{P}_{U'V'}$ überführt, so gibt es zu jedem Geradenpaar p, p' $[p \nmid U, V;$ $p' \nmid U', V']$ eine Kollineation α, so daß*

$$\pi^\varphi = \alpha^{-1}\pi\alpha \qquad \text{für alle } \pi \in \mathfrak{P}_{UV}$$

und $p\alpha = p'$ gilt.

Da die Sätze 1 und 2 zueinander dual sind, genügt es, einen der beiden zu beweisen.

Beweis von Satz 2. Die Abbildung α mit

$$q\alpha = p'\pi^\varphi_{pq} \qquad \text{für } q \nmid U, V \tag{4}$$

ist eine eineindeutige Abbildung der Menge der Geraden q mit $q \nmid U, V$ auf die Menge der Geraden q' mit $q' \nmid U', V'$. Konkurrente Geraden gehen dabei in konkurrente Geraden über, denn seien q, r, s drei voneinander verschiedene Geraden durch einen Punkt P mit $P \neq U, V$, so haben π_{qr}, π_{qs} die gleiche Achse, nämlich (P, V), da sich Gerade und Bildgerade hinsichtlich einer perspektiven Kollineation auf der Achse derselben schneiden und die Achse jedes Elementes von $\mathfrak{P}_{UV}$ durch V geht. Da somit π_{qr}, π_{qs} einer Komponente von $\mathfrak{P}_{UV}$ angehören, gilt dies nach Voraussetzung auch für $\pi^\varphi_{qr}, \pi^\varphi_{qs}$, die also ebenfalls die gleiche Achse haben. Aus $\pi_{pq}\pi_{qr} = \pi_{pr}$ und $\pi_{pq}\pi_{qs} = \pi_{ps}$, also $\pi^\varphi_{pq}\pi^\varphi_{qr} = \pi^\varphi_{pr}$ und $\pi^\varphi_{pq}\pi^\varphi_{qs} = \pi^\varphi_{ps}$ folgt

$$r\alpha = p'\pi^\varphi_{pr} = p'\pi^\varphi_{pq}\pi^\varphi_{qr} = (q\alpha)\pi^\varphi_{qr}$$

$$s\alpha = p'\pi^\varphi_{ps} = p'\pi^\varphi_{pq}\pi^\varphi_{qs} = (q\alpha)\pi^\varphi_{qs}.$$

Mithin schneiden sich nach (1) sowohl $q\alpha$ und $r\alpha$ als auch $q\alpha$ und $s\alpha$ auf der Achse von π^φ_{qr}, woraus sich ergibt, daß $q\alpha, r\alpha, s\alpha$ konkurrent sind.

Hiermit geht ein Geradenbüschel mit Zentrum $P \neq U, V$ in ein Geradenbüschel mit Zentrum $P' \neq U', V'$ über. P' kann als das Bild $P\alpha$ von P hinsichtlich α aufgefaßt werden. Mit dem Bewiesenen ist also gezeigt, daß α eine eineindeutige Abbildung der Menge der Punkte $\neq U, V$ und der Geraden

[12]) Der Kürze halber soll im folgenden der Zusatz „auf der Menge der Punkte P mit $P \nmid u, v$" im Falle der Gruppe $\mathfrak{P}_{uv}$ bzw. „auf der Menge der Geraden g mit $g \nmid U, V$" im Falle der Gruppe $\mathfrak{P}_{UV}$ stets weggelassen werden.

$\nmid U, V$ auf die Menge der Punkte $\neq U', V'$ und der Geraden $\nmid U', V'$ ist, bei der die Inzidenz erhalten bleibt. Dann aber läßt sich α sofort zu einer Kollineation der Ebene erweitern: Die Punkte $P, Q, R, \ldots$ einer Punktreihe a durch U bzw. V gehen in die Punkte einer Punktreihe a' durch U' oder V' über, denn wäre $(P\alpha, Q\alpha) = b \nmid U', V'$, so folgte $b\alpha^{-1} = a \nmid U, V$ im Widerspruch zur Voraussetzung, also gilt $(P\alpha, Q\alpha) \mathrel{I} U'$ oder V' und entsprechend $(P\alpha, R\alpha)$, $(Q\alpha, R\alpha) \mathrel{I} U'$ oder V'. Wäre nun $(P\alpha, Q\alpha) \mathrel{I} U'$, jedoch $(P\alpha, R\alpha) \mathrel{I} V' \neq U'$, so folgte, wenn $P, Q, R \neq U, V$ ist, $(Q\alpha, R\alpha) \nmid U', V'$, im Widerspruch zum eben Bewiesenen. Daher kann man a' als das Bild $a\alpha$ von a hinsichtlich α definieren. Das Geradenbüschel durch U — und entsprechend das Geradenbüschel durch V — geht entweder in das Geradenbüschel durch U' oder in das durch V' über, denn wäre für $a, b, c \mathrel{I} U$, (a, b, c voneinander verschieden) etwa $(a\alpha, b\alpha) = A \neq U', V'$, so folgte $A\alpha^{-1} = a \cdot b \neq U$, im Widerspruch zur Voraussetzung, also gilt $(a\alpha, b\alpha) = U'$ oder V' und entsprechend $(a\alpha, c\alpha)$, $(b\alpha, c\alpha) = U'$ oder V'. Wäre dann $(a\alpha, b\alpha) = U'$, jedoch $(a\alpha, c\alpha) = V' \neq U'$, so folgte für $a, b, c \neq (U, V)$, daß $(b\alpha, c\alpha) \neq U', V'$ gilt, im Widerspruch zum eben Bewiesenen. Somit kann man auch für U und V eindeutig ein Bild hinsichtlich α so definieren, daß die Inzidenz erhalten bleibt. Mithin ist die derart erweiterte Abbildung α eine Kollineation.

Es gilt nun $(q\alpha)\alpha^{-1}\pi\alpha = q\pi\alpha = p'\pi^{\varphi}_{p(q\pi)} = p'\pi^{\varphi}_{pq}\pi^{\varphi}_{q(q\pi)} = p'\pi^{\varphi}_{pq}\pi^{\varphi} = (q\alpha)\pi^{\varphi}$, also — da $\alpha^{-1}\pi\alpha, \pi^{\varphi}$ perspektive Kollineationen $\in \mathfrak{P}_{U'V'}$ sind — nach Lemma 1 $\alpha^{-1}\pi\alpha = \pi^{\varphi}$. Da weiterhin auch $p\alpha = p'\pi^{\varphi}_{pp} = p'$ gilt, ist Satz 2 vollständig bewiesen.

Satz 1 besitzt eine Reihe von wichtigen Spezialfällen und Folgerungen, von denen wir die nachstehenden anführen wollen:

1. *Ist φ ein Automorphismus der einfach transitiven Gruppe $\mathfrak{P}_{vu}$, der jede Komponente in sich überführt, so ist jede nach Satz 1 zu φ gehörende Kollineation α eine perspektive Kollineation mit v als Achse.*

2. *Ist φ ein Automorphismus der einfach transitiven Gruppe $\mathfrak{P}_{vu}$, der jede Komponente wieder in eine Komponente, dabei eine Komponente $\mathfrak{P}_{Zu}[Z \mathrel{I} v]$ elementweise in sich überführt, so ist jede nach Satz 1 zu φ gehörende Kollineation α eine perspektive Kollineation mit $(Z, P) = w$ als Achse und einem Zentrum Z' mit $Z' \mathrel{I} u, v$, wenn $P = P'$ gewählt wird.*

Denn jeder Punkt Q auf w ist Fixpunkt: $Q\alpha = P\pi^{\varphi}_{PQ} = P\pi_{PQ} = Q$, da $\pi_{PQ} \in \mathfrak{P}_{Zu}$. Da ferner u, v Fixgeraden $\neq w$ von α sind, muß das Zentrum von α auf u und v liegen.

3. *Ist die Gruppe $\mathfrak{P}_{vu}$ einfach transitiv, so auch die Gruppe $\mathfrak{P}_{uv}$. Die Elemente $\in \mathfrak{P}_{vu}$ kommutieren mit den Elementen von $\mathfrak{P}_{uv}$.*

Beweis. Zu dem identischen Automorphismus von $\mathfrak{P}_{vu}$ existiert nach Satz 1 für jedes Punktepaar P, P' $[P, P' \nmid u, v]$ eine Kollineation α mit $P\alpha = P'$ und $\pi = \alpha^{-1}\pi\alpha$ für alle $\pi \in \mathfrak{P}_{vu}$. α ist nach 1 perspektive Kollineation mit v als Achse. Das Zentrum von α liegt auf u, denn wegen $\pi^{-1}\alpha\pi = \alpha$ ist das Zentrum von α Fixpunkt für alle $\pi \in \mathfrak{P}_{vu}$ und muß daher mit u inzidieren. Die zweite Aussage von 3 folgt unmittelbar aus $\pi = \alpha^{-1}\pi\alpha$, also $\alpha\pi = \pi\alpha$.

Als Spezialfall ist in 3 enthalten, daß die Gruppe $\mathfrak{P}_{uu}$ abelsch ist.

4. *Ist die Gruppe $\mathfrak{P}_{uu}$ einfach transitiv, so gibt es zu jedem Punkt P mit $P \nmid u$ eine Homologie σ mit P als Zentrum und u als Achse, für die $\sigma^2 = 1$ und $\sigma\tau\sigma = \tau^{-1}$ für alle $\tau \in \mathfrak{P}_{uu}$ gilt. Gibt es eine nichtinvolutorische Translation $\tau \neq 1$ in $\mathfrak{P}_{uu}$, so ist σ involutorisch.*

Beweis. Da nach 3 die Gruppe $\mathfrak{P}_{uu}$ abelsch ist, ist die Abbildung φ mit $\tau^{\varphi} = \tau^{-1}$ ein Automorphismus von $\mathfrak{P}_{uu}$, der jede Komponente in sich überführt. Mithin gehört zu φ nach 1 eine Homologie σ mit u als Achse und P als Zentrum, wobei P beliebig $[P \nmid u]$ gewählt werden kann. Aus $\tau^{\varphi} = \sigma^{-1}\tau\sigma = \tau^{-1}$ folgt $(P\tau)\sigma^2 = P\sigma\tau^{-1}\sigma = P\sigma^2\sigma^{-1}\tau^{-1}\sigma = P\sigma^2\tau = P\tau$, also nach Lemma 1 $\sigma^2 = 1$, da $P\tau \nmid u$ und $P\tau \neq P$ angenommen werden kann.

Mit den Aussagen 1—4 gelten natürlich auf Grund von Satz 2 auch die dualen Aussagen. Diese werden wir im folgenden mit 1′, 2′ usw. zitieren.

Satz 1 und 2 können wir noch den nachstehenden Satz an die Seite stellen, dessen Beweis mutatis mutandis der gleiche ist wie für Satz 2, so daß wir auf eine ausführliche Wiedergabe wohl verzichten dürfen.

Satz 3. *Sind die Gruppen $\mathfrak{P}_{vu}$ und $\mathfrak{P}_{UV}$ einfach transitiv und ist φ ein Isomorphismus von $\mathfrak{P}_{vu}$ und $\mathfrak{P}_{UV}$, der jede Komponente von $\mathfrak{P}_{vu}$ in eine Komponente von $\mathfrak{P}_{UV}$ überführt, so gibt es zu jedem Paar P, p $[P \nmid u, v;\ p \nmid U, V]$ eine Dualität δ, so daß*

$$\pi^{\varphi} = \delta^{-1}\pi\delta \qquad \text{für alle } \pi \in \mathfrak{P}_{vu}$$

und $P\delta = p$ gilt.

§ 3. Einige Beziehungen zwischen Perspektivitäten und perspektiven Kollineationen. Vorbereitungen für § 4

Für das folgende wählen wir drei feste, nicht kollineare Punkte O, U, V mit den Verbindungsgeraden $x = (O, U)$, $y = (O, V)$ und $u = (U, V)$. Ferner sei eine Gerade $\bar{z}$ mit $\bar{z} \mid U$, jedoch $\bar{z} \neq x, u$ fest gewählt. Dann führen wir zur Abkürzung nachstehende Bezeichnungen ein:

$$\sigma_A = \pi_{xAu};$$
$$\varrho_{Az} = \pi_{xAz} \quad \text{für } A \mid u \text{ und } z \mid U;$$
$$\varphi_A = \pi_{xAy} \quad \text{für } A \mid u.$$

Speziell soll darüber hinaus noch $\varrho_{Vz} = \varrho_z$ und $\varrho_{A\bar{z}} = \varrho_A$ sein. Ist nun X ein Punkt auf x, so bezeichnen wir den Punkt $X\varrho_V$ mit X', den Punkt $X\varphi_Z$ [für fest gewählten Punkt $Z \neq U, V$ auf u] mit $\tilde{X}$ und den Punkt $X\sigma_{O'}$ mit $\bar{X}$.

Die Abbildungen

$$\bar{\tau}_A = \sigma_{O'}\sigma_{A'}^{-1} \quad \text{und} \quad \bar{\tau}_A^{*} = \varrho_V\varrho_A^{-1} \qquad \text{für } A \mid x \text{ mit } A \neq U$$

sind projektive[13]) Abbildungen der Geraden x auf sich, für die

$$O\bar{\tau}_A = A,\ U\bar{\tau}_A = U; \quad O\bar{\tau}_A^{*} = A,\ U\bar{\tau}_A^{*} = U$$

[13]) Unter einer projektiven Abbildung einer Geraden a auf eine Gerade b versteht man eine Abbildung von a auf b, die sich als Produkt von Perspektivitäten schreiben läßt.

gilt. Zu jedem $A \neq U$ auf x gibt es eineindeutig eine Abbildung $\bar{\tau}_A$ und eineindeutig eine Abbildung $\bar{\tau}_A^*$. Die Menge der Abbildungen $\bar{\tau}_A$ bezeichnen wir mit $\mathfrak{T}$, die Menge der Abbildungen $\bar{\tau}_A^*$ mit $\mathfrak{T}^*$. Offenbar gilt nun

$$X\bar{\tau}_A = A\bar{\tau}_X^* \qquad \textit{für } A, X \neq U \textit{ auf } x. \tag{5}$$

Die Menge $\mathfrak{T}$ (und entsprechend die Menge $\mathfrak{T}^*$) ist genau einfach transitiv auf der Menge der Punkte $\neq U$ auf x, denn seien A, B zwei beliebige Punkte $\neq U$ auf x und $C = B\bar{\tau}_A^{*-1}$, so gilt wegen (5) $A\bar{\tau}_C = B$.

Wir definieren nun eine Verknüpfung für die Elemente aus $\mathfrak{T}$, die nicht notwendig mit dem Hintereinanderausführen der Abbildungen übereinzustimmen braucht:

$$\bar{\tau}_A \circ \bar{\tau}_B = \bar{\tau}_C \textit{ mit } C = O\bar{\tau}_A\bar{\tau}_B = A\bar{\tau}_B{}^{14}). \tag{6}$$

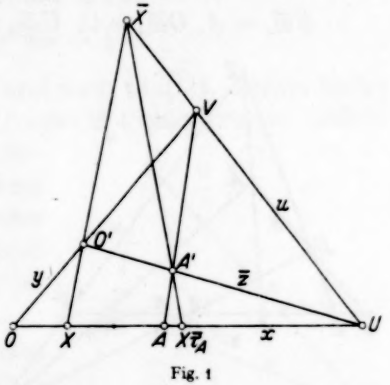


Fig. 1

Es gilt dann

Lemma 5. *Die Menge $\mathfrak{T}$ bildet hinsichtlich der Verknüpfung* (6) *ein Loop*[15]).

Beweis. Offenbar ist $\bar{\tau}_O$ die identische Abbildung der Geraden x auf sich, und daher gilt $\bar{\tau}_A \circ \bar{\tau}_O = \bar{\tau}_A$ und $\bar{\tau}_O \circ \bar{\tau}_A = \bar{\tau}_A$. Ferner sind die Gleichungen $\bar{\tau}_A \circ \bar{\tau}_X = \bar{\tau}_B$ und $\bar{\tau}_Y \circ \bar{\tau}_A = \bar{\tau}_B$ für vorgegebene $\bar{\tau}_A$, $\bar{\tau}_B$ stets eindeutig lösbar, denn sei $X = B\bar{\tau}_A^{*-1}$, also nach (5) $A\bar{\tau}_X = B$, und $Y = B\bar{\tau}_A^{-1}$, so gilt $O\bar{\tau}_A\bar{\tau}_X = A\bar{\tau}_X = B$ und $O\bar{\tau}_Y\bar{\tau}_A = Y\bar{\tau}_A = B$ und damit die Behauptung.

Das Loop $\mathfrak{T}$ ist dann und nur dann eine Gruppe, wenn die Verknüpfung (6) mit dem Hintereinanderausführen der Abbildungen $\bar{\tau}_A$ übereinstimmt, wenn also gilt

$$\bar{\tau}_A \circ \bar{\tau}_B = \bar{\tau}_A\bar{\tau}_B = \bar{\tau}_{A\bar{\tau}_B}, \tag{7}$$

denn das Hintereinanderausführen von Abbildungen genügt stets dem assoziativen Gesetz, und andererseits folgt aus

$$X(\bar{\tau}_A \circ \bar{\tau}_B) = O\left(\bar{\tau}_X \circ (\bar{\tau}_A \circ \bar{\tau}_B)\right) = O\left((\bar{\tau}_X \circ \bar{\tau}_A) \circ \bar{\tau}_B\right) = (X\bar{\tau}_A)\bar{\tau}_B = X\bar{\tau}_A\bar{\tau}_B$$

für alle $X \neq U$ auf x Gl. (7).

Ist $\mathfrak{T}$ eine Gruppe, so gilt

$$\bar{\tau}_A^*\bar{\tau}_B = \bar{\tau}_B\bar{\tau}_A^*, \tag{8}$$

da nach (5) und (7) $(X\bar{\tau}_A^*)\bar{\tau}_B = A\bar{\tau}_X\bar{\tau}_B = A\bar{\tau}_{X\bar{\tau}_B} = (X\bar{\tau}_B)\bar{\tau}_A^*$ für alle $X \neq U$ auf x ist.

[14]) D.h., C ist das Bild des Punktes O hinsichtlich der Abbildung $\bar{\tau}_A\bar{\tau}_B$ (die ja nicht notwendig in $\mathfrak{T}$ enthalten ist).

[15]) Ein Loop ist ein algebraisches System $\mathfrak{S}$, in dem eine Verknüpfung $\circ$ erklärt ist, für die folgende Regeln gelten: (a) Zu $a, b \in \mathfrak{S}$ gibt es genau ein $c \in \mathfrak{S}$ mit $a \circ b = c$. (b) Zu $b, c \in \mathfrak{S}$ gibt es genau ein x mit $x \circ b = c$. (c) Zu $a, c \in \mathfrak{S}$ gibt es genau ein y mit $a \circ y = c$. (d) Es gibt genau ein „neutrales" Element n mit $a \circ n = a$ und $n \circ b = n$ für alle $a, b \in \mathfrak{S}$.

Gl. (8) können wir auch die folgende Form geben:

$$(8') \qquad \varrho_V^{-1}\bar{\tau}_B\varrho_V = \varrho_{V'}^{-1}\bar{\tau}_B\varrho_{V'} \qquad \textit{mit bel. } V' \neq U \textit{ auf } u$$

(man setze nämlich $\bar{A} = V'$).

Es ist unmittelbar klar, daß die Menge $\mathfrak{T}^*$ eine Gruppe bildet, wenn $\mathfrak{T}$ eine Gruppe ist, und umgekehrt. Wir betrachten weiterhin die Abbildungen

$$\bar{\omega}_A = \sigma_{O'}\sigma_{\bar{A}}^{-1} \quad \textit{und} \quad \bar{\omega}_A^* = \varphi_Z\varphi_{\bar{A}}^{-1} \quad \textit{für } A \neq O, U \textit{ auf } x$$

der Geraden x auf sich. Ist $O'\varphi_Z^{-1} = E$, so gilt für sie

$$E\bar{\omega}_A = A,\ O\bar{\omega}_A = O,\ U\bar{\omega}_A = U;\ E\bar{\omega}_A^* = A,\ O\bar{\omega}_A^* = O,\ U\bar{\omega}_A^* = U.$$

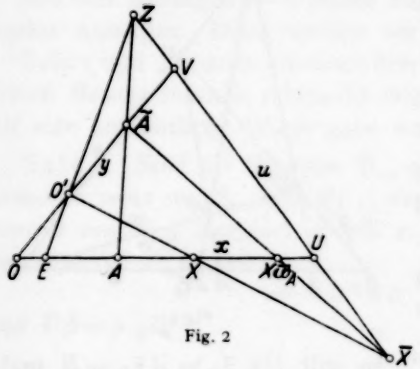


Fig. 2

Zu jedem $A \neq O, U$ auf x gibt es eineindeutig eine Abbildung $\bar{\omega}_A$ und eineindeutig eine Abbildung $\bar{\omega}_A^*$. Die Menge der Abbildungen $\bar{\omega}_A$ bezeichnen wir mit $\mathfrak{H}$, die der Abbildungen $\bar{\omega}_A^*$ mit $\mathfrak{H}^*$. $\mathfrak{H}$ und $\mathfrak{H}^*$ sind jede genau einfach transitiv auf der Menge der Punkte $\neq O, U$ auf x. Es gilt auch hier

$$(9) \qquad X\bar{\omega}_A = A\bar{\omega}_X^* \textit{ für } A, X \neq O, U \overset{\epsilon}{} \textit{ auf } x.$$

Wir definieren für die Abbildungen $\bar{\omega}_A$ analog wie für die Abbildungen $\bar{\tau}_A$ eine Verknüpfung, die nicht notwendig das Hintereinanderausführen der Abbildungen $\bar{\omega}_A$ darstellt:

$$(10) \qquad \bar{\omega}_A \cdot \bar{\omega}_B = \bar{\omega}_C \quad \textit{mit} \quad C = E\bar{\omega}_A\bar{\omega}_B = A\bar{\omega}_B.$$

Es gilt dann entsprechend wie für die Menge $\mathfrak{T}$ das

Lemma 6. *Die Menge $\mathfrak{H}$ bildet hinsichtlich der Verknüpfung* (10) *ein Loop.*

Das neutrale Element ist dabei die identische Abbildung $\bar{\omega}_E$ von x auf sich.

Weiter schließt man wie oben, daß $\mathfrak{H}$ genau dann eine Gruppe bildet, wenn die Verknüpfung (10) dem Hintereinanderausführen der Abbildungen $\bar{\omega}_A$ entspricht. Ferner ergibt sich auch hier im Falle, daß $\mathfrak{H}$ eine Gruppe ist, die Gleichung

$$(11) \qquad \bar{\omega}_A\bar{\omega}_B^* = \bar{\omega}_B^*\bar{\omega}_A \qquad \textit{für } A, B \neq O, U \textit{ auf } x.$$

Für das Zusammensetzen der Abbildungen $\in \mathfrak{T}$ und $\in \mathfrak{H}$ gilt folgende, später verwendete Tatsache: Dann und nur dann gibt es zu jedem Punkt P mit $P \not\mathrel{I} u$, x Elemente $\bar{\tau}_X, \bar{\omega}_Y$ mit

$$(12) \qquad \sigma_{O'}\sigma_P^{-1} = \bar{\omega}_Y\bar{\tau}_X,$$

wenn die Abbildungen $\bar{\tau}_A$ unabhängig von der Wahl von $\bar{z}$ durch U sind. Denn sind die Abbildungen $\bar{\tau}_A$ unabhängig von der Wahl von $\bar{z}$, so gilt für $z = (P, U)$, $Y = (z, y)\varphi_Z^{-1}$, $X = P\varrho_z^{-1}$ ersichtlich $\bar{\omega}_Y = \sigma_{O'}\sigma_{\bar{Y}}^{-1}$ und $\bar{\tau}_X = \sigma_{\bar{Y}}\sigma_P^{-1}$,

also (12), und umgekehrt folgt aus (12) $\bar{\tau}_X = \sigma_{\bar{Y}} \sigma_P^{-1}$, also die genannte Unabhängigkeit.

Wir wollen nun untersuchen, welche weitergehenden Aussagen sich über die Abbildungen $\bar{\tau}_A$, $\bar{\omega}_A$ ergeben, wenn wir die Existenz spezieller perspektiver Kollineationen fordern. Als erstes behandeln wir das Axiom

P4. *Die Gruppe* $\mathfrak{P}_{U\,u}$ *ist linear transitiv.*

Es gibt dann zu jedem $A \neq U$ auf x eine und nur eine Translation $\tau \in \mathfrak{P}_{U\,u}$ mit $O\tau = A$. Diese bezeichnen wir mit τ_A. Man hat dann

$$X \tau_A = X \bar{\tau}_A \qquad \textit{für alle } X \mathrel{I} x, \tag{13}$$

$$X \tau_A \tau_B = X \bar{\tau}_A \circ \bar{\tau}_B \qquad \textit{für alle } X \mathrel{I} x. \tag{14}$$

Letzteres folgt aus $O \tau_A \tau_B = A \tau_B = O \bar{\tau}_A \circ \bar{\tau}_B$ und nach Gl. (13). Mithin bilden die Abbildungen $\bar{\tau}_A$ bei Voraussetzung von Axiom P4 eine Gruppe. Außerdem aber sieht man ohne weiteres, daß die Abbildungen $\bar{\tau}_A$ dann unabhängig sind von der Wahl der Geraden $\bar{z}$ durch U, denn wie auch immer man die Gerade $\bar{z} \neq u, x$ gewählt hat, stets muß Gl. (13) gelten, wodurch $\bar{\tau}_A$ als Abbildung der Geraden x auf sich eindeutig festgelegt ist.

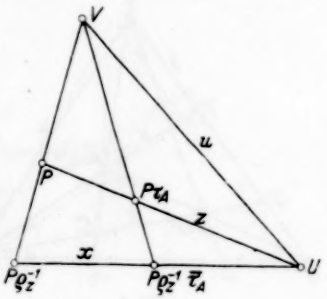


Fig. 3

Wir wollen nun zeigen, daß es umgekehrt zu jeder Abbildung $\bar{\tau}_A \in \mathfrak{T}$ eine Translation $\tau_A \in \mathfrak{P}_{U\,u}$ gibt, in die sie eingebettet werden kann, wenn die Abbildungen $\bar{\tau}_A$ Gruppeneigenschaft und die eben genannte Unabhängigkeit besitzen. Wir definieren dazu τ_A vorerst als eineindeutige Abbildung der Menge der Punkte auf sich durch

$$P \tau_A = P \varrho_z^{-1} \bar{\tau}_A \varrho_z \qquad \textit{für } P \not\mathrel{I} u \textit{ und } z = (P, U)$$

$$P \tau_A = P \qquad \textit{für } P \mathrel{I} u.$$

Zu zeigen ist, daß τ_A drei kollineare Punkte P_1, P_2, P_3 in drei kollineare Punkte überführt. Dies ist jedoch klar, wenn P_1, P_2, P_3 auf einer Geraden durch V, dem Zentrum der Perspektivität ϱ_z liegen. Ferner ergibt sich aus Gl. (8'), die aus der Gruppeneigenschaft von $\mathfrak{T}$ folgt, wegen der Unabhängigkeit der Abbildungen $\in \mathfrak{T}$ von der Wahl von $\bar{z}$

$$P \varrho_{V z}^{-1} \bar{\tau}_A \varrho_{V z} = P \varrho_{V' z}^{-1} \bar{\tau}_A \varrho_{V' z} \qquad \textit{mit bel. } V' \neq U \textit{ auf } u,$$

d.h. die Definition von τ_A ist unabhängig von der Wahl des Zentrums V' der Perspektivität von x auf z, wenn nur V' auf u liegt. Hieraus folgt sofort die Behauptung auch für Punkte P_i $(i = 1, 2, 3)$, die auf einer Geraden durch einen beliebigen Punkt $V' \neq U$ auf u liegen. Da endlich die Punkte auf einer Geraden z mit $z \mathrel{I} U$ offenbar in Punkte auf z abgebildet werden, ist τ_A eine Kollineation, die u punktweise und U geradenweise festläßt, mithin eine Translation $\in \mathfrak{P}_{U\,u}$, die natürlich mit $\bar{\tau}_A$ auf x übereinstimmt. Damit ist bewiesen der

Satz 4. *In einer projektiven Ebene gilt Axiom* P4 *genau dann, wenn folgende Bedingungen erfüllt sind:*

1. *Das Loop* $\mathfrak{T}$ *ist eine Gruppe.*
2. *Die Elemente* $\in \mathfrak{T}$ *sind unabhängig von der Wahl von* $\bar{z}$ *durch* U.

Die Abbildungen

$$\bar{\tau}'_P = \sigma_{O'}^{-1}\sigma_{P'} \quad \textit{für jedes } P \neq U \textit{ auf } u \textit{ mit } P' = P\pi_{uO\bar{z}}$$

sind projektive Abbildungen der Geraden u auf sich, für die

$$V\bar{\tau}'_P = P, \quad U\bar{\tau}'_P = U$$

gilt. Sie entsprechen genau den Abbildungen $\bar{\tau}_A$ der Geraden x auf sich. Die Menge der Abbildungen $\bar{\tau}'_P$ bezeichnen wir mit $\mathfrak{T}'$. Zu jedem $\bar{\tau}'_P$ gibt es ein und nur ein Element $\bar{\tau}_A \in \mathfrak{T}$ mit

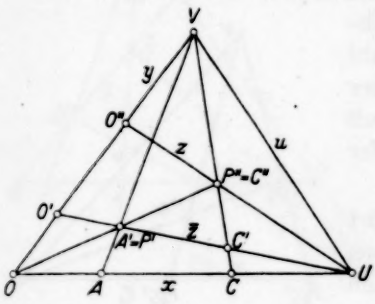


Fig. 4

$$\bar{\tau}'^{-1}_P = \sigma_{O'}^{-1}\bar{\tau}_A\sigma_{O'}, \tag{15}$$

denn setzt man $P'\varrho_V^{-1} = A$, so gilt für dies A Gl. (15). Mithin bildet $\mathfrak{T}'$ eine Gruppe (hinsichtlich des Hintereinanderausführens der Abbildungen), wenn $\mathfrak{T}$ eine Gruppe ist, und umgekehrt.

Sind nun die Abbildungen $\bar{\tau}'_P$ von der Wahl von $\bar{z}$ unabhängig, d.h. ist z eine beliebige Gerade $\neq x$, u durch U und gilt

$$\bar{\tau}'_P = \sigma_{O''}^{-1}\sigma_{P''} \tag{16}$$

mit $O'' = V\pi_{uOz}$ und $P'' = P\pi_{uOz}$, so gibt es dann und nur dann eine Abbildung $\bar{\tau}_C \in \mathfrak{T}$ mit

$$\bar{\tau}'^{-1}_P = \sigma_{O''}^{-1}\bar{\tau}_C\sigma_{O''}, \tag{17}$$

wenn auch die Abbildungen $\bar{\tau}_A$ von $\bar{z}$ unabhängig sind, wenn also auch

$$\bar{\tau}_C = \sigma_{O''}\sigma_{C''}^{-1} \quad \textit{mit } C'' = C\varrho_z \tag{18}$$

gilt (man setze nämlich $C = P''\varrho_z^{-1}$, also $P'' = C''$).

Aus (15) und (17) folgt nun sofort

$$\sigma_{O''}\sigma_{O'}^{-1}\bar{\tau}_A\sigma_{O'}\sigma_{O''}^{-1} = \bar{\tau}_C, \quad \textit{also} \quad \bar{\omega}_B^{-1}\bar{\tau}_A\bar{\omega}_B = \bar{\tau}_C$$

mit $B = O''\varphi_Z^{-1}$, also $O'' = \tilde{B}$. Da nun $O\bar{\omega}_B^{-1}\bar{\tau}_A\bar{\omega}_B = A\bar{\omega}_B$ ist, gilt

$$\bar{\omega}_B^{-1}\bar{\tau}_A\bar{\omega}_B = \bar{\tau}_{A\bar{\omega}_B} \quad \textit{für alle } \bar{\tau}_A \in \mathfrak{T},\ \bar{\omega}_B \in \mathfrak{H}. \tag{19}$$

Umgekehrt folgt aus (19) und (18) wegen (15) sofort (16), d.h. wir haben bewiesen:

Lemma 7. *Gelten von den drei Forderungen*

1. *Die Elemente von* $\mathfrak{T}$ *sind von der Wahl von* $\bar{z}$ *unabhängig.*
2. *Die Elemente von* $\mathfrak{T}'$ *sind von der Wahl von* $\bar{z}$ *unabhängig.*

3. *Es gilt* $\bar{\omega}_B^{-1}\bar{\tau}_A\bar{\omega}_B = \bar{\tau}_{A\bar{\omega}_B}$ *für alle* $\bar{\tau}_A \in \mathfrak{T}$, $\bar{\omega}_B \in \mathfrak{H}$. *zwei, so auch die dritte.*

Betrachten wir nun das Axiom

P5. *Die Gruppe* $\mathfrak{P}_{U_x}$ *ist linear transitiv.*

Dann kann man wie für Satz 4 zeigen:

Satz 5. *In einer projektiven Ebene gilt Axiom* P5 *genau dann, wenn die Abbildungen* $\bar{\tau}'_P$ *eine Gruppe bilden und unabhängig von der Wahl von* $\bar{z}$ *sind.*

Zusammen mit Lemma 7 ergibt dieser Satz sofort den weiteren:

Satz 6. *Unter Voraussetzung von* P4 *gilt Axiom* P5 *in einer projektiven Ebene genau dann, wenn Gl.* (19) *richtig ist.*

Denn gilt P4, so ist nach Satz 4 $\mathfrak{T}$ eine Gruppe, also auch $\mathfrak{T}'$. Ferner sind dann die Abbildungen $\bar{\tau}_A$ unabhängig von der Wahl von $\bar{z}$ und somit muß nach Lemma 7 und Satz 5 Axiom P5 gelten, wenn Gl. (19) richtig ist, und umgekehrt folgt aus P5 nach Satz 5 und Lemma 7, daß (19) gültig ist.

Wir betrachten nun das Axiom

P6: *Die Gruppe* $\mathfrak{P}_{V_u}$ *ist linear transitiv.*

Dann gilt ein Satz 6 entsprechender

Satz 7. *Unter Voraussetzung von* P4 *gilt* P6 *genau dann, wenn die Gleichung gilt:*

$$(20)\quad \begin{cases} \bar{\omega}_B^{*-1}\bar{\tau}_A\bar{\omega}_B^* = \bar{\tau}_{A\bar{\omega}_B^*} \\ \text{für alle } \bar{\tau}_A \in \mathfrak{T},\ \bar{\omega}_B^* \in \mathfrak{H}^*. \end{cases}$$

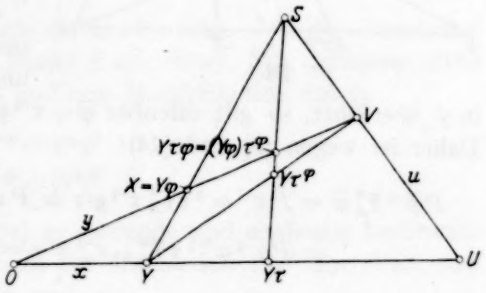


Fig. 5

Beweis. (a) Sei $\mathfrak{P}_{V_u}$ linear transitiv und $\tau \in \mathfrak{P}_{U_u}$ und φ eine Perspektivität π_{xSy} mit $S \mathrel{I} u$, dann bezeichnen wir die Translation $\in \mathfrak{P}_{V_u}$, die O in $(O\tau)\varphi$ überführt mit τ^φ. Wegen $(O\tau)\tau^{-1}\tau^\varphi = O\tau^\varphi = (O\tau)\varphi$ ist $\tau^{-1}\tau^\varphi \in \mathfrak{P}_{S_u}$. Ist nun Y ein Punkt $\neq U$ auf x, so muß daher $Y\tau^\varphi = (Y\tau)\tau^{-1}\tau^\varphi$ auf $(Y\tau, S)$ liegen. Da $Y\tau^\varphi$ natürlich auch auf (Y, V) liegt, muß $Y\tau^\varphi = (Y, V)\cdot(Y\tau, S)$ und daher $(Y\varphi)\tau^\varphi = (Y\tau)\varphi$ sein. Setzt man nun $X = Y\varphi$, so erhält man sofort

$$(21)\qquad X\varphi^{-1}\tau\varphi = X\tau^\varphi \quad \text{für bel. } \tau \in \mathfrak{P}_{U_u} \text{ und } X \mathrel{I} y.$$

Die Abbildung $\tau \leftrightarrow \tau^\varphi$ ist also ein Isomorphismus von $\mathfrak{P}_{U_u}$ auf $\mathfrak{P}_{V_u}$. Da nun $\bar{\omega}_B^* = \varphi_Z\varphi_B^{-1}$ gilt, folgt aus (21) zunächst $X\varphi_Z^{-1}\tau_A\varphi_Z = X\tau_{A\varphi_Z}$ für alle X auf y und weiter aus (21) $X\varphi_B^{-1}\tau_{A\varphi_Z\bar{\varphi}_B^{-1}}\varphi_B = X\tau_{A\varphi_Z}$ für alle X auf y, also $X\varphi_B\tau_{A\varphi_Z}\varphi_B^{-1} = X\tau_{A\bar{\omega}_B^*}$ mit $X \mathrel{I} x$, woraus sich

$$(22)\quad X\bar{\omega}_B^{*-1}\tau_A\bar{\omega}_B^* = X\tau_{A\bar{\omega}_B^*} \quad \text{für alle } X \text{ auf } x \text{ und } A\ B \neq U, O \text{ auf } x,$$

und mit (13) die Gl. (20) ergibt.

(b) Sei Gl. (20) gültig. Dann konstruieren wir eine Translation $\tau' \in \mathfrak{P}_{V_u}$, die O in einen beliebigen Punkt A' auf y überführt, indem wir definieren

$$P\tau' = P\bar{\varphi}^{-1}\bar{\tau}_A^*\bar{\varphi} \quad \text{für } P \not\mathrel{I} u$$
$$P\tau' = P \quad \text{für } P \mathrel{I} u,$$

wobei $A = A'\varphi_Z^{-1}$, $\bar{\varphi} = \pi_{x Z \bar{y}}$ mit $\bar{y} = (P, V)$ ist. Wie im Beweis zu Satz 4 ergibt sich die Tatsache, daß τ' Kollineation ist, aus der „Unabhängigkeitsaussage":

$$P\bar{\varphi}^{-1}\bar{\tau}_A^*\bar{\varphi} = P\bar{\varphi}'^{-1}\bar{\tau}_{A\bar{\omega}^*}^*\bar{\varphi}', \tag{23}$$

wo $\bar{\varphi}' = \pi_{x Z' \bar{y}}$ mit $Z' \neq U, V$ auf u und $\bar{\omega}^* = \varphi_Z\varphi_{Z'}^{-1}$ ist. Diese erschließt man wie folgt: Aus (20) ergibt sich, wenn zur Abkürzung $\varphi_Z = \varphi$ und $\varphi_{Z'} = \varphi'$ gesetzt wird,

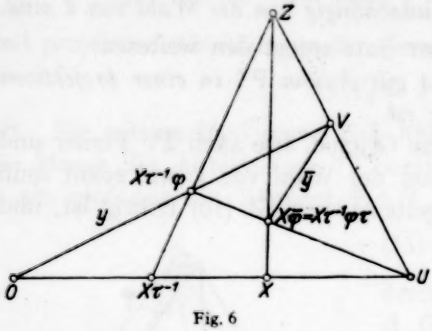


Fig. 6

$$X\varphi^{-1}\tau_S\varphi = X\varphi'^{-1}\tau_{S\bar{\omega}^*}\varphi'$$
$$\text{für} \quad S \mathrel{I} x; \quad X \mathrel{I} y; \quad S \mathrel{\not I} u,$$

und daher — für $S = P\varphi^{-1}$ und $X = A\varphi$ — $(A\varphi)\varphi^{-1}\tau_{P\varphi^{-1}}\varphi = (A\varphi)\varphi'^{-1}\tau_{P\varphi'^{-1}}\varphi'$, also

$$P\varphi^{-1}\bar{\tau}_A^*\varphi = P\varphi'^{-1}\bar{\tau}_{A\bar{\omega}^*}^*\varphi', \tag{24}$$

mithin Gl. (23) für $P \mathrel{I} y$. Ist nun $P \mathrel{I} \bar{y}$ und τ die Translation $\in \mathfrak{P}_{U_u}$, die y in $\bar{y}$ überführt, so gilt offenbar $\bar{\varphi} = \tau^{-1}\varphi\tau$ (s. Fig. 6) und auch $\bar{\varphi}' = \tau^{-1}\varphi'\tau$. Daher ist wegen (8) und (24):

$$P\bar{\varphi}^{-1}\bar{\tau}_A^*\bar{\varphi} = P\tau^{-1}\varphi^{-1}\tau\bar{\tau}_A^*\tau^{-1}\varphi\tau = P\tau^{-1}\varphi^{-1}\bar{\tau}_A^*\varphi\tau = P\tau^{-1}\varphi'^{-1}\bar{\tau}_{A\bar{\omega}^*}^*\varphi'\tau$$
$$= P\tau^{-1}\varphi'^{-1}\tau\bar{\tau}_{A\bar{\omega}^*}^*\tau^{-1}\varphi'\tau = P\bar{\varphi}'^{-1}\bar{\tau}_{A\bar{\omega}^*}^*\bar{\varphi}',$$

also gilt Gl. (23).

Daß $\tau' \in \mathfrak{P}_{V_u}$ mit $O\tau' = A'$ gilt, ist nach Definition klar, und damit ist Satz 7 bewiesen.

Zum Schluß dieses Paragraphen wollen wir den Beweis, daß die Gültigkeit von Axiom P4 und P5 das Bestehen der Gl. (19) zur Folge hat, noch einmal entsprechend dem Teil (a) des Beweises von Satz 7 führen, da die dabei verwendeten Schlüsse später noch einmal benutzt werden.

Seien also die Gruppen $\mathfrak{P}_{U_u}$ und $\mathfrak{P}_{U_x}$ linear transitiv, und σ eine Perspektivität π_{xAu}, dann bezeichnen wir mit τ^σ für $\tau \in \mathfrak{P}_{U_u}$ die Translation $\in \mathfrak{P}_{U_x}$, die den Punkt $O\sigma$ in den Punkt $(O\tau)\sigma$ überführt. Wir behaupten dann:

(i) *Ist* $(A, U) = z$, *so ist* $\tau\tau^\sigma \in \mathfrak{P}_{U_x}$ *für jedes* $\tau \in \mathfrak{P}_{U_u}$.

Beweis. Da $(O\tau)\tau^\sigma = O\tau$ und $(O\sigma)\tau\tau^\sigma = (O\sigma)\tau^\sigma = (O\tau)\sigma$ ist, geht $(O, O\sigma) = g$ bei $\tau\tau^\sigma$ in $(O\tau, (O\tau)\sigma) = h$ über. Der Schnittpunkt A der Geraden g mit ihrer Bildgeraden h liegt aber auf der Achse von $\tau\tau^\sigma$, also hat $\tau\tau^\sigma$ z als Achse.

(ii) *Ist* $\tau \in \mathfrak{P}_{U_u}$, *so gilt* $X\sigma^{-1}\tau\sigma = X\tau^\sigma$ *für alle* X *auf* u.

Beweis. Sei $Y = X\sigma^{-1}$, so schneiden sich $(Y, Y\sigma)$ und $\big(Y\tau, (Y\tau)\sigma\big)$ in A. Wegen $Y\tau\tau^\sigma = Y\tau$ und (i) ist dann $(Y, Y\sigma)\tau\tau^\sigma = \big(Y\tau, (Y\tau)\sigma\big)$, woraus $(Y\sigma)\tau\tau^\sigma = Y\sigma\tau^\sigma = (Y\tau)\sigma$, mithin $X\tau^\sigma = X\sigma^{-1}\tau\sigma$ folgt.

Zweimaliges Anwenden von (ii) liefert dann Gl. (19).

§ 4. Einführung von Koordinaten

Für das Folgende setzen wir voraus, daß die Gruppe $\mathfrak{P}_{U_u}$ linear transitiv ist.

Ist $\overline{\omega}$ ein Element von $\mathfrak{H}$ und τ eine Translation $\in \mathfrak{P}_{U_u}$, so gibt es wegen der linearen Transitivität von $\mathfrak{P}_{U_u}$ eine Translation $\in \mathfrak{P}_{U_u}$, die den Punkt O in den Punkt $(O\tau)\overline{\omega}$ überführt. Wir bezeichnen diese mit $\tau^{\overline{\omega}}$. Jedes Element aus $\mathfrak{H}$ induziert somit eine eineindeutige Abbildung $\tau \leftrightarrow \tau^{\overline{\omega}}$ der Gruppe $\mathfrak{P}_{U_u}$ auf sich, wobei verschiedene Elemente $\in \mathfrak{H}$ auch verschiedene Abbildungen induzieren, da aus $(O\tau)\overline{\omega} = (O\tau)\overline{\omega}'$ sofort $\overline{\omega} = \overline{\omega}'$ folgt. Die Menge dieser Abbildungen, samt dem Endomorphismus 0, der alle Elemente $\in \mathfrak{P}_{U_u}$ in das Einselement abbildet, bezeichnen wir mit $\mathfrak{A}$, deren Elemente mit $\alpha, \beta, \gamma, \ldots$.

Da das Loop $\mathfrak{H}$ genau einfach transitiv auf der Menge der Punkte $\neq O, U$ auf x ist, gilt:

(*) Sind $\tau \neq 1$ und τ' zwei Translationen $\in \mathfrak{P}_{U_u}$, so gibt es ein und nur ein $\alpha \in \mathfrak{A}$ mit $\tau^\alpha = \tau'$.

Im folgenden bezeichnet τ, wenn nicht anders bemerkt, stets die Translation $\in \mathfrak{P}_{U_u}$, welche den Punkt O in den Punkt E überführt. Wir definieren dann für die Elemente $\in \mathfrak{A}$ eine Addition und eine Multiplikation durch

$$(**) \qquad \begin{cases} \tau^{\alpha+\beta} = \tau^\alpha \tau^\beta \\ \tau^{\alpha\cdot\beta} = (\tau^\alpha)^\beta. \end{cases}$$

Dies bedeutet: $\alpha+\beta$ ist das nach (*) existierende und eindeutig bestimmte Element $\in \mathfrak{A}$, das die Translation τ in die Translation $\tau^\alpha\tau^\beta$ überführt, und $\alpha \cdot \beta$ ist das nach (*) existierende und eindeutig bestimmte Element $\in \mathfrak{A}$, das τ in $(\tau^\alpha)^\beta$ überführt[16]).

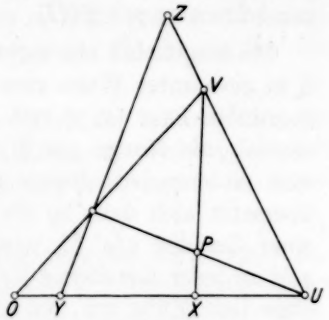


Fig. 7

Es ist klar, daß $\mathfrak{A}$ hinsichtlich der eingeführten Addition eine zu $\mathfrak{P}_{U_u}$ isomorphe Gruppe und hinsichtlich der Multiplikation — ohne 0 natürlich — ein zu $\mathfrak{H}$ isomorphes Loop bildet. Ferner gilt $\alpha \cdot 0 = 0 \cdot \alpha$ für alle $\alpha \in \mathfrak{A}$.

Wir wollen nun für die Punkte der projektiven Ebene (affine) Koordinaten einführen. Dazu zeichnen wir die Gerade u als uneigentliche Gerade aus und wählen die Geraden x und y aus § 3 als Koordinatenachsen. Sei dann P ein Punkt mit $P \nmid u$ und $z = (P, U)$, ferner $X = P\varrho_z^{-1}$ und $Y = (z, y)\varphi_Z^{-1}$, dann gibt es wegen (*) Elemente $\xi, \eta \in \mathfrak{A}$ mit $\tau^\xi = \tau_X$ und $\tau^\eta = \tau_Y$. ξ und η heißen die Koordinaten von P. Damit ist jedem Punkt P mit $P \nmid u$ eineindeutig ein geordnetes Paar (ξ, η) von Koordinaten zugeordnet.

Eine Gerade g mit $g \nmid U$ läßt sich dann durch eine lineare Gleichung beschreiben, denn sei $g \cdot x = B$ und $(g, u)\sigma_O^{-1} = A$, so gilt wegen (12) $A\overline{\omega}_Y\tau_X = B$

[16]) Man beachte hierbei, daß das so definierte Multiplizieren von Abbildungen $\in \mathfrak{A}$ im allgemeinen nicht mit dem Hintereinanderausführen derselben übereinstimmt.

für alle Punkte auf g. Hieraus erhält man $O\tau_A\bar{\omega}_Y\tau_X = O\tau_B$, woraus sich $\tau_A^{\bar{\omega}_Y}\tau_X = \tau_B$, also, wenn α, β nach (*) durch $\tau_A = \tau^\alpha$ und $\tau_B = \tau^\beta$ bestimmt werden, $(\tau^\alpha)^\eta\tau^\xi = \tau^\beta$ oder $\alpha\cdot\eta+\xi=\beta$ ergibt. Mithin erfüllen die Koordinaten ξ, η aller Punkte P auf g die Gleichung

$$\alpha\cdot\eta+\xi=\beta. \tag{25}$$

Die Koordinaten ξ, η der Punkte auf einer Geraden g mit $g \mathrel{I} U$ erfüllen natürlich die Gleichung $\eta=\beta$.

Da in der projektiven Ebene zwei verschiedene Geraden einen und nur einen Schnittpunkt und zwei verschiedene Punkte eine und nur eine Verbindungsgerade besitzen, folgen aus den Geradengleichungen (25) die Auflösbarkeitsbedingungen

A1. *Die Gleichung* $\alpha\cdot\eta-\beta\cdot\eta=\gamma$ *für* $\alpha, \beta, \gamma \in \mathfrak{A}$ *und* $\alpha\neq\beta$ *besitzt eine eindeutige Lösung* η *in* $\mathfrak{A}$.

A2. *Die Gleichung* $-\xi\cdot\alpha+\xi\cdot\beta=\gamma$ *für* $\alpha, \beta, \gamma \in \mathfrak{A}$ *und* $\alpha\neq\beta$ *besitzt eine eindeutige Lösung* ξ *in* $\mathfrak{A}$.

Ein algebraisches System $\mathfrak{C}$, in dem eine Addition und eine Multiplikation erklärt ist und das hinsichtlich der Addition eine Gruppe (mit 0 als neutralem Element), hinsichtlich der Multiplikation (ohne 0) aber ein Loop bildet und in dem $\alpha\cdot 0 = 0\cdot\alpha$ für alle $\alpha\in\mathfrak{C}$ und die Auflösbarkeitsforderungen A1 und A2 gelten, bezeichnen wir als ein cartesisches Zahlensystem[17]). Mit dem oben Bewiesenen ist also gezeigt, daß die Menge $\mathfrak{A}$, die wir aus der projektiven Ebene Π gewonnen haben, ein cartesisches Zahlensystem ist. Wir bezeichnen es mit $\mathfrak{C}(\Pi)$.

Ist umgekehrt ein cartesisches Zahlensystem $\mathfrak{C}$ gegeben, so läßt sich aus $\mathfrak{C}$ in gewohnter Weise eine affine Ebene gewinnen. Man faßt nämlich jedes geordnete Paar (x, y) mit $x, y\in\mathfrak{C}$ als Punkt auf und definiert als Geraden einmal jede Menge $g(a, b)$ von Punkten (x, y) mit $ay+x=b$ $[a, b\in\mathfrak{C}]$ und zum anderen jede Menge $g(b)$ von Punkten (x, y) mit $y=b$. Die Inzidenz übersetzt sich dann in die $\in$-Relation. Ferner heißt eine Gerade $g(a, b)$ zu einer Geraden $g(a', b')$ parallel, wenn $a=a'$ ist, und außerdem jede Gerade $g(b)$ zu jeder Geraden $g(b')$ parallel. Die affinen Inzidenzaxiome folgen dann ohne jede Mühe aus den algebraischen Eigenschaften von $\mathfrak{C}$.

Fassen wir dann jedes Büschel paralleler Geraden als Punkte einer uneigentlichen Geraden auf, so entsteht aus der affinen Ebene eine projektive Ebene, die wir die projektive Koordinatenebene über $\mathfrak{C}$ nennen und mit $\Pi(\mathfrak{C})$ bezeichnen wollen.

Erklärt man nun in $\Pi(\mathfrak{C})$ die Abbildungen $\bar{\tau}_A, \bar{\omega}_A$ in bezug auf das Koordinatensystem von $\Pi(\mathfrak{C})$, also mit x und y als den Koordinatenachsen, u als der uneigentlichen Geraden und endlich mit E und O' als den Punkten $(1, 0)$ und $(0, 1)$, so gilt

$$O\bar{\tau}_A\bar{\tau}_B = (a+b, 0);\ O\bar{\omega}_A\bar{\omega}_B = (ab, 0) \quad \textit{für} \quad A=(a, 0) \ \textit{und} \ B=(b, 0), \tag{26}$$

[17]) In [13] cartesische Gruppe genannt.

denn es ist $(A, O') = (A, \bar{A}) = g(a, a)$ und $B' = (b, 1)$ und $\bar{B} = (0, b)$, also $(\bar{A}, B') = g(a, a+b)$ und $(\bar{A}, \bar{B}) = g(a, ab)$, woraus $A\,\bar{\tau}_B = (a+b, 0)$ und $A\,\bar{\omega}_B = (ab, 0)$, mithin (26) folgt.

Aus (26) folgt insbesondere, daß die Abbildungen $\bar{\tau}_A$ eine Gruppe bilden. Die Abbildungen $\bar{\tau}_A$ sind ferner unabhängig von der Wahl von $\bar{z}$ durch U, denn ist $P = (x, y)$ ein Punkt auf einer Geraden $g = g(a, b)$ $[g \nmid U]$, so gilt $ay + x = b$, also nach (26) $A\,\bar{\omega}_Y\,\bar{\tau}_X = B$, wenn $X = (x, 0)$ und $Y = (y, 0)$ ist. Da andererseits auch $A\,\sigma_{O'}\,\sigma_P^{-1} = B$ gilt, hat man $A\,\bar{\omega}_Y\,\bar{\tau}_X = A\,\sigma_0\,\sigma_P^{-1}$. Dies gilt jedoch für alle Geraden g durch P $[g \nmid U]$, also alle $A \neq U$ auf x, woraus $\bar{\omega}_Y\,\bar{\tau}_X = \sigma_{O'}\,\sigma_P^{-1}$, mithin nach S. 342 die Behauptung folgt.

Nach Satz 4 ist damit die Gruppe $\mathfrak{P}_{U\,u}$ linear transitiv. Die für $\mathfrak{P}_{U\,u}$ gebildete Menge $\mathfrak{A}$ muß dann nach dem oben Bewiesenen ein cartesisches Zahlensystem bilden, welches nach (26) isomorph zu $\mathfrak{C}$ ist. Wir können daher zusammengenommen als Satz formulieren:

Satz 8. (a) *Zu jedem cartesischen Zahlensystem $\mathfrak{C}$ gibt es eine projektive Koordinatenebene $\Pi(\mathfrak{C})$ über $\mathfrak{C}$, in der die Gruppe $\mathfrak{P}_{U\,u}$ linear transitiv ist, wenn u die uneigentliche Gerade und U der Schnittpunkt der x-Achse mit u ist.*

(b) *Jeder projektiven Ebene Π, in der es eine Gerade u und einen Punkt U auf u derart gibt, daß die Gruppe $\mathfrak{P}_{U\,u}$ linear transitiv ist, kann in der oben beschriebenen Weise ein cartesisches Zahlensystem $\mathfrak{C}(\Pi)$ — bezogen auf U, u — zugeordnet werden. Π ist dann darstellbar als projektive Koordinatenebene über $\mathfrak{C}(\Pi)$.*

(c) *Ist $\mathfrak{C}$ ein cartesisches Zahlensystem, $\Pi(\mathfrak{C})$ die nach* (a) *zu $\mathfrak{C}$ gehörige projektive Koordinatenebene, $\mathfrak{C}(\Pi(\mathfrak{C}))$ das nach* (b) *zu $\Pi(\mathfrak{C})$ gehörige cartesische Zahlensystem, so gilt $\mathfrak{C}(\Pi(\mathfrak{C})) \cong \mathfrak{C}$, wenn U, u in* (a) *und* (b) *die gleichen Elemente sind.*

Die Aussagen (a) und (b) in Satz 8 wollen wir in der kurzen Form zusammenfassen: *Eine projektive Ebene ist genau dann als projektive Koordinatenebene über einem cartesischen Zahlensystem darstellbar, wenn in ihr Axiom* P 4 *gilt.*

Ersetzt man in den Entwicklungen dieses Paragraphen die Gruppe $\mathfrak{P}_{U\,u}$ durch die Menge $\mathfrak{T}$ und fordert die Unabhängigkeit der Elemente $\in \mathfrak{T}$ von der Wahl von $\bar{z}$, so sieht man, daß sich fast alles entsprechend definieren und beweisen läßt. Nur ist das algebraische System $\mathfrak{S}$, zu dem die analog definierte Menge $\mathfrak{A}$ durch Einführen der Verknüpfungen (**) wird, im allgemeinen hinsichtlich der Addition nicht mehr assoziativ, sondern bildet ein zu $\mathfrak{T}$ isomorphes Loop. Alle anderen Rechengesetze eines cartesischen Zahlensystems gelten auch in $\mathfrak{S}$ (wobei man allerdings die Auflösbarkeitsbedingungen in der ursprünglichen, direkt aus (25) folgenden Form stehen lassen muß, da das Assoziativgesetz der Addition in $\mathfrak{S}$ nicht notwendig gilt). Da sich über $\mathfrak{S}$ genau wie über $\mathfrak{C}$ eine projektive Koordinatenebene erklären läßt, hat man daher den Satz, daß sich eine projektive Ebene genau dann als projektive Koordinatenebene über einem solchen algebraischen System $\mathfrak{S}$ darstellen läßt, wenn die Abbildungen $\bar{\tau}_A$ unabhängig von der Wahl von $\bar{z}$ sind.

§ 5. Stufenaufbau der Begründung der projektiven Geometrie

Satz 8 kann man noch eine Reihe entsprechender Sätze an die Seite stellen, die ebenfalls projektive Ebenen, in denen geeignete Zusatzaxiome gelten, als projektive Koordinatenebenen über gewissen algebraischen Systemen kennzeichnen.

Setzen wir z.B. noch außer der Gültigkeit von P4 voraus, daß auch P5 erfüllt ist, so gilt nach Gl. (19): $\tau^{\bar\omega} = \bar\omega^{-1}\tau\bar\omega$ *für alle* $\tau \in \mathfrak{P}_{U_u}$. Mithin ist die Abbildung $\tau \leftrightarrow \tau^{\bar\omega}$ *für alle* $\tau \in \mathfrak{P}_{U_u}$ ein Automorphismus von $\mathfrak{P}_{U_u}$, also sind alle Elemente von $\mathfrak{A}$ Endomorphismen der Gruppe $\mathfrak{P}_{U_u}$. Dann aber läßt sich sofort die Gültigkeit des Linksdistributivgesetzes herleiten:

$$\tau^{(\alpha+\beta)\cdot\gamma} = (\tau^\alpha\tau^\beta)^\gamma = (\tau^\alpha)^\gamma(\tau^\beta)^\gamma = \tau^{\alpha\cdot\gamma}\tau^{\beta\cdot\gamma} = \tau^{\alpha\cdot\gamma+\beta\cdot\gamma}.$$

Umgekehrt, gilt in $\mathfrak{A}$ das Linksdistributivgesetz, so ist jedes Element $\alpha \in \mathfrak{A}$ Endomorphismus, denn sind τ_1, τ_2 zwei Translationen $\in \mathfrak{P}_{U_u}$, so gibt es nach (*) Elemente $\alpha, \beta \in \mathfrak{A}$ mit $\tau^\beta = \tau_1$ und $\tau^\gamma = \tau_2$, und es gilt

$$(\tau_1\tau_2)^\alpha = (\tau^\beta\tau^\gamma)^\alpha = (\tau^{\beta+\gamma})^\alpha = \tau^{\beta\cdot\alpha+\gamma\cdot\alpha} = \tau^{\beta\cdot\alpha}\tau^{\gamma\cdot\alpha} = (\tau^\beta)^\alpha(\tau^\gamma)^\alpha = \tau_1^\alpha\tau_2^\beta.$$

Damit ist die Abbildung $\tau \leftrightarrow \tau^{\bar\omega}$ [für alle $\tau \in \mathfrak{P}_{U_u}$] für jedes $\bar\omega \in \mathfrak{H}$ ein Automorphismus von $\mathfrak{P}_{U_u}$, für den wegen $X\tau\bar\omega = O\tau_X\tau\bar\omega = O(\tau_X\tau)^{\bar\omega} = O\tau_X^{\bar\omega}\tau^{\bar\omega} = X\bar\omega\tau^{\bar\omega}$ für alle X auf x gilt

$$\tau^{\bar\omega} = \bar\omega^{-1}\tau\bar\omega \qquad \textit{für alle } \tau \in \mathfrak{P}_{U_u}.$$

Dies bedeutet aber, daß Gl. (19) gültig, also nach Satz 6, daß die Gruppe $\mathfrak{P}_{U_x}$ linear transitiv ist. Bezeichnet man eine cartesische Gruppe, in der ein Distributivgesetz gilt, als ein Veblen-Wedderburn-System[18]) mit Links- bzw. Rechtsdistributivgesetz, so haben wir nach Satz 8 bewiesen:

Satz 9. *Eine projektive Ebene ist genau dann als projektive Koordinatenebene über einem Veblen-Wedderburn-System mit Linksdistributivgesetz darstellbar, wenn in ihr die Axiome* P4 *und* P5 *gelten.*

Welche Folgerungen für den Koordinatenbereich $\mathfrak{A}$ kann man nun ziehen, wenn statt P5 zusätzlich P6 gilt? Bezeichnen wir wie oben mit $\tau^{\bar\omega^*}$ für $\bar\omega^* \in \mathfrak{H}^*$ die Translation, die den Punkt O in den Punkt $(O\tau)\bar\omega^*$ überführt, so besitzt die Menge $\mathfrak{A}^*$ der eineindeutigen Abbildungen $\tau \leftrightarrow \tau^{\bar\omega^*}$ *für alle* $\tau \in \mathfrak{P}_{U_u}$ der Gruppe $\mathfrak{P}_{U_u}$ auf sich ebenfalls die Eigenschaft (*), wenn wir noch den Endomorphismus 0 hinzunehmen, denn auch die Menge $\mathfrak{H}^*$ ist genau einfach transitiv auf der Menge der Punkte $\neq O, U$ auf x.

Zu jedem $\alpha \in \mathfrak{A}$ gibt es daher genau ein $\alpha' \in \mathfrak{A}^*$ mit $\tau^\alpha = \tau^{\alpha'}$. Ferner hat $E\bar\omega_A\bar\omega_B = A\bar\omega_B = B\bar\omega_A^* = E\omega_B^*\bar\omega_A^*$ die Gleichung $(\tau^\alpha)^\beta = (\tau^{\beta'})^{\alpha'}$ zur Folge, wenn $\alpha, \beta \in \mathfrak{A}$; $\alpha', \beta' \in \mathfrak{A}^*$ und $\tau^\alpha = \tau^{\alpha'}$; $\tau^\beta = \tau^{\beta'}$ gilt.

Die Gültigkeit von Axiom P6 bedeutet nun wegen Gl. (20), daß $\tau^{\bar\omega^*} = \bar\omega^{*-1}\tau\bar\omega^*$ *für alle* $\tau \in \mathfrak{P}_{U_u}$ gilt, also jedes Element $\neq 0$ aus $\mathfrak{A}^*$ Automorphismus

[18]) In [13] Quasikörper genannt.

der Gruppe $\mathfrak{P}_{U_u}$ ist. Daraus erschließen wir nun sofort das Rechtsdistributivgesetz:

$$\tau^{\alpha\cdot(\beta+\gamma)} = (\tau^{\alpha})^{\beta+\gamma} = (\tau^{(\beta+\gamma)'})^{\alpha'} = (\tau^{\beta+\gamma})^{\alpha'} = (\tau^{\beta})^{\alpha'}(\tau^{\gamma})^{\alpha'} = (\tau^{\beta'})^{\alpha'}(\tau^{\gamma'})^{\alpha'}$$
$$= (\tau^{\alpha})^{\beta}(\tau^{\alpha})^{\gamma} = \tau^{\alpha\cdot\beta+\alpha\cdot\gamma}.$$

Hierbei sind $\alpha', \beta', (\alpha+\beta)'$ die Elemente aus $\mathfrak{A}^*$, für die $\tau^{\alpha} = \tau^{\alpha'}$, $\tau^{\beta} = \tau^{\beta}$ und $\tau^{(\alpha+\beta)'} = \tau^{(\alpha+\beta)}$ gilt.

Umgekehrt, gilt in $\mathfrak{A}$ das Rechtsdistributivgesetz, so müssen die Elemente $\in \mathfrak{A}^*$ Endomorphismen sein: Sind τ_1, τ_2 zwei Translationen $\in \mathfrak{P}_{U_u}$, so gibt es Elemente β, γ mit $\tau^{\beta} = \tau_1$ und $\tau^{\gamma} = \tau_2$. Daher gilt

$$(\tau_1\tau_2)^{\alpha'} = (\tau^{\beta}\tau^{\gamma})^{\alpha'} = (\tau^{\beta+\gamma})^{\alpha'} = (\tau^{(\beta+\gamma)'})^{\alpha'} = (\tau^{\alpha})^{\beta+\gamma}$$
$$= \tau^{\alpha\cdot\beta+\alpha\cdot\gamma} = (\tau^{\alpha})^{\beta}(\tau^{\alpha})^{\gamma} = (\tau^{\beta'})^{\alpha'}(\tau^{\gamma'})^{\alpha'} = \tau_1^{\alpha'}\tau_2^{\alpha'}.$$

Entsprechend wie oben folgt dann, daß $\tau^{\bar{\omega}^*} = \bar{\omega}^{*-1}\tau\bar{\omega}^*$ *für alle* $\tau \in \mathfrak{P}_{U_u}$ ist, also Gl. (20) richtig ist, woraus man nach Satz 7 auf die lineare Transitivität von $\mathfrak{P}_{V_u}$ schließt. Damit ist gezeigt:

Satz 10. *Eine projektive Ebene ist genau dann als projektive Koordinatenebene über einem Veblen-Wedderburn-System mit Rechtsdistributivgesetz darstellbar, wenn in ihr die Axiome* P4 *und* P6 *gelten*[19]).

Aus (*) folgt im Fall, daß P4 und P5 bzw. P4 und P6 erfüllt sind, daß alle Translationen $\neq 1$ aus $\mathfrak{P}_{UU}$ bzw. aus $\mathfrak{P}_{u_u}$ die gleiche Ordnung haben müssen, da dann alle Elemente $\neq 0$ von $\mathfrak{A}$ bzw. $\mathfrak{A}^*$ Automorphismen sind und außerdem jede Komponente von $\mathfrak{P}_{UU}$ bzw. $\mathfrak{P}_{u_u}$ isomorph zu $\mathfrak{P}_{U_u}$ ist [s. (21) und (ii) aus § 3]. Die Ordnung der Translationen $\neq 1$ ist — falls endlich — gleich einer Primzahl p, denn jede Potenz einer Translation $\tau' \neq 1$ hat die gleiche Ordnung wie τ'.

Darüber hinaus gilt, wenn wir mit α_s den nach (*) existierenden und eindeutig bestimmten Endomorphismus $\in \mathfrak{A}$ bezeichnen, der τ (das festgewählte Element $\in \mathfrak{P}_{U_u}$) in τ^s (s ganze Zahl) überführt, der

Satz 11. *Gilt Axiom* P4 *und* P5, *so bildet die Menge* $\mathfrak{R}$ *der Endomorphismen* $\alpha_s\alpha_t^{-1}$ *von* $\mathfrak{P}_{U_u}$ *mit* $t \neq 0$ *in* $\mathfrak{A}$ *einen Körper, der isomorph zum Primkörper der Charakteristik* p *ist, wenn die Translationen* $\in \mathfrak{P}_{U_u}$ *die Ordnung* p *haben, oder isomorph zum Primkörper der Charakteristik* 0 *ist, wenn die Translationen* $\in \mathfrak{P}_{U_u}$ *nicht endliche Ordnung haben.*

Beweis. Zunächst gilt

$$\tau^{\alpha_s+\alpha_t} = \tau^{\alpha_s}\tau^{\alpha_t} = \tau^s\tau^t = \tau^{s+t} = \tau^{\alpha_{s+t}}$$
$$\tau^{\alpha_s\cdot\alpha_t} = (\tau^{\alpha_s})^{\alpha_t} = (\tau^s)^{\alpha_t} = (\tau^{\alpha_t})^s = (\tau^t)^s = \tau^{s\cdot t} = \tau^{\alpha_{st}},$$

[19]) Wir erwähnen noch, daß die Gültigkeit von P4 und P6 im Falle, daß es wenigstens eine nichtinvolutorische Translation $\neq 1$ gibt, gleichwertig mit der Forderung ist, daß es zu jedem Punkt P mit $P \not\mathrel{I} u$ eine involutorische Homologie σ mit P als Zentrum und u als Achse gibt. Dies ergibt sich sofort aus Satz 1, Folgerung 4 und der leicht zu beweisenden Tatsache, daß das Produkt zweier involutorischer Homologien mit gleicher Achse eine Translation ist. Vgl. [12], Teil I und [13], S. 210ff.

wobei außer den Definitionen von Summe und Produkt lediglich benutzt wird, daß α_t ein Endomorphismus ist. Hieraus folgt, daß die Endomorphismen $\alpha_s \in \mathfrak{A}$ einen Ring $\mathfrak{R}$ bilden, der isomorph zum Restklassenring mod p oder isomorph zum Ring der ganzen Zahlen ist, je nachdem die Ordnung der Translationen endlich (nämlich gleich p) oder nicht endlich ist. Da im ersten Fall jedes Element $\alpha_t \in \mathfrak{R}\ [t \neq 0]$ ein Inverses in $\mathfrak{R}$ besitzt, ferner aber das Inverse α_t^{-1} von α_t in $\mathfrak{R}$ eindeutig bestimmt ist, ist $\mathfrak{K} = \mathfrak{R}$ und unser Satz in diesem Fall bewiesen. Der allgemeine Beweis verläuft wie folgt:

(a) *Ist* $\alpha_s \in \mathfrak{R}$ *und sind* α, β *beliebige Elemente von* $\mathfrak{A}$, *so gilt* $(\alpha_s \cdot \beta) \cdot \gamma = \alpha_s \cdot (\beta \cdot \gamma)$. Denn es ist $\tau^{(\alpha_s \cdot \beta) \cdot \gamma} = \big((\tau^{\alpha_s})^{\beta}\big)^{\gamma} = \big((\tau^{s})^{\beta}\big)^{\gamma} = \big((\tau^{\beta})^{s}\big)^{\gamma} = \big((\tau^{\beta})^{\gamma}\big)^{s} = (\tau^{\beta \cdot \gamma})^{s} = (\tau^{s})^{\beta \cdot \gamma} = \tau^{\alpha_s \cdot (\beta \cdot \gamma)}$.

(b) *Ist* $\alpha_s \in \mathfrak{R}\ [s \neq 0]$ *und* β *ein beliebiges Element* $\in \mathfrak{A}$, *so gilt* $\alpha_s^{-1}(\alpha_s \beta) = \beta$. Sei nämlich $\alpha_s^{-1}(\alpha_s \beta) = \gamma$, dann ist nach (a): $\alpha_s[\alpha_s^{-1}(\alpha_s \beta)] = (\alpha_s \alpha_s^{-1})(\alpha_s \beta) = \alpha_s \beta$, also $\alpha_s \gamma = \alpha_s \beta$. Da $\mathfrak{A}$ ein Loop hinsichtlich der Multiplikation ist, folgt hieraus $\gamma = \beta$ und damit die Behauptung.

Nach (a) und (b) folgt aus $\alpha_t \alpha_q = \alpha_{tq}$ mit $t, q \neq 0$, also $(\alpha_t \alpha_q)\, \alpha_{tq}^{-1} = 1$, daß $\alpha_t^{-1} = \alpha_q \alpha_{tq}^{-1}$ und entsprechend $\alpha_q^{-1} = \alpha_t \alpha_{tq}^{-1}$ gilt. Somit ergibt sich nach (a) und wegen der Gültigkeit des einen Distributivgesetzes in $\mathfrak{A}$ [für $t, q \neq 0$]:

$$(27)\quad \begin{cases} \alpha_s \alpha_t^{-1} + \alpha_p \alpha_q^{-1} = \alpha_s(\alpha_q \alpha_{tq}^{-1}) + \alpha_p(\alpha_t \alpha_{tq}^{-1}) = (\alpha_s \alpha_q)\, \alpha_{tq}^{-1} + (\alpha_p \alpha_t)\, \alpha_{tq}^{-1} \\ \qquad = (\alpha_s \alpha_q + \alpha_p \alpha_t)\, \alpha_{tq}^{-1} = (\alpha_s \alpha_q + \alpha_p \alpha_t)\, (\alpha_t \alpha_q)^{-1}. \end{cases}$$

Ferner hat man nach (a), (b) und wegen $\alpha_p \alpha_q^{-1} = (\alpha_p \alpha_t)\, \alpha_{tq}^{-1} = \alpha_t(\alpha_p \alpha_{tq}^{-1})$:

$$(28)\quad \begin{cases} (\alpha_s \alpha_t^{-1})\, (\alpha_p \alpha_q^{-1}) = \alpha_s[\alpha_t^{-1}(\alpha_p \alpha_q^{-1})] = \alpha_s[\alpha_t^{-1}\{\alpha_t(\alpha_p \alpha_{tq}^{-1})\}] \\ \qquad = \alpha_s(\alpha_p \alpha_{tq}^{-1}) = (\alpha_s \alpha_p)\, \alpha_{tq}^{-1} = (\alpha_s \alpha_p)\, (\alpha_t \alpha_q)^{-1}. \end{cases}$$

(27) und (28) entnimmt man aber, daß $\mathfrak{K}$ isomorph zum Quotientenkörper von $\mathfrak{R}$ ist, womit alles bewiesen ist.

Satz 11 hat zur Folge, daß sich jedes Veblen-Wedderburn-System als Vektorraum über dem Primkörper der Charakteristik p bzw. 0 auffassen läßt.

Wir setzen nun den Stufenaufbau fort und betrachten das Axiom

P7. a) *Die Gruppe* $\mathfrak{P}_{O\,u}$ *ist linear transitiv.* b) *Die Gruppe* $\mathfrak{P}_{U\,u}$ *besteht nicht nur aus dem Einselement.*

Aus Axiom P7 folgt nach Lemma 4, daß die Gruppe $\mathfrak{P}_{x\,u}$ einfach transitiv ist, da die nach P7b existierende Translation $\neq 1$ aus $\mathfrak{P}_{U\,u}$ eine Kollineation ist, die u aber nicht O festläßt. Insbesondere also ist $\mathfrak{P}_{U\,u}$ linear transitiv. Weiter gilt:

Lemma 8. *Unter Voraussetzung von Axiom* P7 *gibt es zu zwei Translationen* $\tau, \tau' \neq 1$ *aus* $\mathfrak{P}_{U\,u}$ *eine und nur eine Homologie* $\omega \in \mathfrak{P}_{O\,u}$ *mit* $\omega^{-1} \tau \omega = \tau'$.

Beweis. Wegen der linearen Transitivität der Gruppe $\mathfrak{P}_{O\,u}$ gibt es ein $\omega \in \mathfrak{P}_{O\,u}$ mit $(O\tau)\omega = (O\tau')$, denn $O\tau, O\tau'$ sind von O und U verschieden. Für dieses ω gilt dann $O\omega^{-1}\tau\omega = O\tau\omega = O\tau'$ und daher, da sicher $\omega^{-1}\tau\omega \in \mathfrak{P}_{U\,u}$ ist, nach Lemma 1 $\omega^{-1}\tau\omega = \tau'$. Gäbe es noch eine weitere Homologie $\omega' \in \mathfrak{P}_{O\,u}$

mit $\omega'^{-1}\tau\omega'=\omega^{-1}\tau\omega$, so folgte aus $(O\tau)\omega'=O\omega'^{-1}\tau\omega'=O\omega^{-1}\tau\omega=(O\tau)\omega$ nach Lemma 1 $\omega'=\omega$.

Die Gruppe $\mathfrak{A}'$ der durch

$$\tau^{\psi}=\omega^{-1}\tau\omega \qquad \tau\in\mathfrak{P}_{U_u},\ \omega\in\mathfrak{P}_{O_u}$$

definierten Automorphismen ψ der Gruppe $\mathfrak{P}_{U_u}$ samt dem Endomorphismus 0 hat also die Eigenschaft (*). Ist nun τ eine beliebige, aber fest gewählte Translation $\neq 1$ aus $\mathfrak{P}_{U_u}$ und definiert man mit diesem τ eine Addition und Multiplikation in $\mathfrak{A}'$ gemäß (**), so stimmt die Multiplikation mit dem Hintereinanderausführen der Endomorphismen überein, denn $\alpha\cdot\beta$ für $\alpha,\beta\in\mathfrak{A}'$ ist der Endomorphismus, der die (festgewählte) Translation τ in $(\tau^{\alpha})^{\beta}$ überführt. Dieses tut aber auch der Endomorphismus $\alpha\beta$, der durch Hintereinanderausführen von α,β entsteht, und da $\alpha\beta\in\mathfrak{A}'$ ist, muß wegen (*) $\alpha\cdot\beta=\alpha\beta$ gelten. Also bildet $\mathfrak{A}'$ nicht nur hinsichtlich der Addition, sondern auch hinsichtlich der Multiplikation (ohne 0 natürlich) eine Gruppe. Diese ist selbstverständlich isomorph zu $\mathfrak{P}_{O_u}$. Da aus der Endomorphismeneigenschaft der Elemente $\in\mathfrak{A}'$ das eine Distributivgesetz folgt, ferner auch die Auflösbarkeitsbedingung A2 gelten muß, bildet $\mathfrak{A}'$ hinsichtlich der eingeführten Verknüpfungen ein Veblen-Wedderburn-System mit Linksdistributivgesetz, in dem die Multiplikation assoziativ ist. Ein solches System nennt man einen Fastkörper (mit Linksdistributivgesetz)[20]. Der Koordinatenbereich, der zu einer projektiven Ebene gehört, in der Axiom P7 gilt, ist also ein Fastkörper mit Linksdistributivgesetz.

Dieses Ergebnis, das wir hier direkt ableiteten, kann man natürlich auch leicht aus den vorher durchgeführten Entwicklungen entnehmen, denn bezeichnet ω_A die nach P7 existierende Homologie $\in\mathfrak{P}_{O_u}$, die den Punkt E in den Punkt A überführt, so gilt analog zu (8) und (9):

$$X\omega_A=X\overline{\omega}_A \qquad \textit{für alle } X \textit{ auf } x, \tag{29}$$

$$X\omega_A\omega_B=X\overline{\omega}_A\cdot\overline{\omega}_B \qquad \textit{für alle } X \textit{ auf } x. \tag{30}$$

Wegen (29) ist dann aber $\mathfrak{A}'=\mathfrak{A}$ und wegen (30) die Multiplikation in $\mathfrak{A}$ assoziativ.

Umgekehrt, ist ein Fastkörper $\mathfrak{F}$ gegeben, so gelten in der projektiven Koordinatenebene $\Pi(\mathfrak{F})$ über $\mathfrak{F}$ Axiom P4 und P5. Da $\mathfrak{F}$ (ohne Nullelement) nach (26) hinsichtlich der Multiplikation isomorph zu $\mathfrak{H}$ ist, bilden die Abbildungen $\overline{\omega}_A$ in $\Pi(\mathfrak{F})$ eine Gruppe. Wir zeigen nun, daß sich hieraus die lineare Transitivität der Gruppe $\mathfrak{P}_{O_u}$ ergibt.

Da die Gruppe $\mathfrak{P}_{UU}$ direktes Produkt der Gruppen $\mathfrak{P}_{U_u}$ und $\mathfrak{P}_{U_x}$ ist, läßt sich jedes Element $\tau\in\mathfrak{P}_{UU}$ eindeutig in der Form $\tau_1\tau_2$ mit $\tau_1\in\mathfrak{P}_{U_x}$ und $\tau_2\in\mathfrak{P}_{U_u}$ darstellen. Die Abbildung

$$\tau^{\psi}=\tau_1\tau_2^{\overline{\omega}} \qquad \textit{mit festem } \overline{\omega}\in\mathfrak{H}$$

[20]) In [13] assoziativer Quasikörper genannt. Die Definition eines Fastkörpers nach Zassenhaus [Über endliche Fastkörper. Abh. math. Seminar Univ. Hamburg **11**, 123—147 (1931)] enthält nicht die Auflösbarkeitsbedingungen A1 und A2.

ist dann ein Automorphismus von $\mathfrak{P}_{UU}$, da $\tau_2 \leftrightarrow \tau_2^{\bar{\omega}}$ ein Automorphismus von $\mathfrak{P}_{Uu}$ und die Gruppe $\mathfrak{P}_{UU}$ abelsch ist. Der Automorphismus ψ führt jede Komponente $\mathfrak{P}_{Uz}$ von $\mathfrak{P}_{UU}$ wieder in eine Komponente über, denn ist $z = u$ oder $z = x$, so ist die Behauptung richtig. Ist $z \neq x, u$ und $(z, y) = \widetilde{B}$, dann gilt offenbar für alle $\tau = \tau_1 \tau_2 \in \mathfrak{P}_{Uz}$: $\tau_2 = \tau_1^{\sigma_{\widetilde{B}}^{-1}}$ (Bezeichnungen analog wie in § 3). Hieraus folgt $\tau^{\psi} = \tau_1 \tau_1^{\sigma_{\widetilde{B}}^{-1}\bar{\omega}}$. Wegen der Gruppeneigenschaft der Abbildungen $\bar{\omega}_A$ gibt es dann ein $\widetilde{C}$ mit $\sigma_{O'} \sigma_{\widetilde{B}}^{-1} \bar{\omega} = \sigma_{O'} \sigma_{\widetilde{C}}^{-1}$, also mit $\sigma_{\widetilde{B}}^{-1} \bar{\omega} = \sigma_{\widetilde{C}}^{-1}$. Also gilt $\tau^{\psi} = \tau_1 \tau_1^{\sigma_{\widetilde{C}}^{-1}}$, woraus nach (i) am Schluß von § 2 die Behauptung folgt.

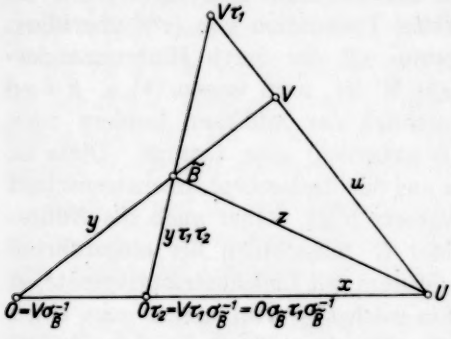


Fig. 8

Nach Satz 2, Folgerung 2′ gehört zu dem Automorphismus ψ eine Homologie mit u als Achse und O als Zentrum, wenn $p = p' = y$ gewählt wird, denn ψ führt $\mathfrak{P}_{Ux}$ elementweise in sich. Wegen der Transitivität der Abbildungen $\in \mathfrak{H}$ muß dann auch die Gruppe $\mathfrak{P}_{Ou}$ linear transitiv sein. Es gilt also der

Satz 12. *Eine projektive Ebene ist genau dann als Koordinatenebene über einem Fastkörper mit Linksdistributivgesetz darstellbar, wenn in ihr Axiom* P 7 *gilt.*

Satz 10 und Satz 12 ergeben bei Beachtung von Lemma 4 den

Satz 13. *Eine projektive Ebene ist genau dann als projektive Koordinatenebene über einem Schiefkörper darstellbar, wenn in ihr die Gruppe $\mathfrak{P}_{Ou}$ linear transitiv ist und die Gruppen $\mathfrak{P}_{Uu}$ und $\mathfrak{P}_{Vu}$ nicht nur aus dem Einselement bestehen.*

Gilt bei Voraussetzung von Axiom P 7 noch zusätzlich, daß die Gruppe $\mathfrak{P}_{Ou}$ abelsch ist, so ist klar, daß die Multiplikation in $\mathfrak{A}$ kommutativ ist, also $\mathfrak{A}$ ein kommutativer Körper ist. Umgekehrt folgt natürlich aus der Kommutativität in $\mathfrak{A}$, daß $\mathfrak{P}_{Ou}$ abelsch ist. Man kann also die projektiven Ebene über einem kommutativen Körper durch Axiom P 7 mit der zusätzlichen Forderung kennzeichnen, daß $\mathfrak{P}_{Ou}$ abelsch ist. Diese zusätzliche Forderung kann aber durch eine Existenzforderung ersetzt werden (s. Satz 14)[21]), nämlich das Axiom

P 8. *Es gibt eine perspektive Dualität δ mit O als Zentrum und u als Achse.*

Dem Beweis schicken wir folgendes Lemma voraus:

Lemma 9. *Ist die Gruppe $\mathfrak{P}_{vu}$ einfach transitiv und eine Komponente $\mathfrak{P}_{Ou}$ $[O \not\mathrel{I} u,\ O \mathrel{I} v]$ abelsch, so ist jede Komponente von $\mathfrak{P}_{vu}$ und $\mathfrak{P}_{uv}$ abelsch.*

Beweis. Die Behauptung ist klar für alle Komponenten von $\mathfrak{P}_{vu}$, da diese — außer $\mathfrak{P}_{Uu}$ mit $U \mathrel{I} u$ — durch innere Automorphismen von $\mathfrak{P}_{vu}$ auf $\mathfrak{P}_{Ou}$ abgebildet werden können. $\mathfrak{P}_{Uu}$ aber ist abelsch, weil nach Satz 1,

[21]) Vgl. BAER [4].

Folgerung 3 $\mathfrak{P}_{uv}$ einfach transitiv, also $\mathfrak{P}_{Uv}$ linear transitiv, mithin $\mathfrak{P}_{UU}$ einfach transitiv ist.

Seien nun π', π'' Elemente einer Komponente der Gruppe $\mathfrak{P}_{uv}$ und P ein beliebiger Punkt mit $P \nmid u, v$. Dann gibt es Elemente $\omega', \omega'' \in \mathfrak{P}_{vu}$ mit $P\omega' = P\pi'$ und $P\omega'' = P\pi''$. Da die Elemente $\in \mathfrak{P}_{uv}$ mit denen aus $\mathfrak{P}_{vu}$ kommutieren (vgl. Satz 1, Folgerung 3) und nach dem eben Bewiesenen auch ω', ω'' kommutieren, ergibt sich:

$$(P\pi')\pi'' = (P\omega')\pi'' = (P\pi'')\omega' = (P\omega'')\omega' = (P\omega')\omega'' = (P\pi')\omega''$$
$$= (P\omega'')\pi' = (P\pi'')\pi', \text{ also } \pi'\pi'' = \pi''\pi'.$$

Wir zeigen nun, daß bei Voraussetzung von Axiom P7 die Existenz einer perspektiven Dualität $\delta \in \mathfrak{D}_{Ou}$ die Kommutativität der Homologien $\in \mathfrak{P}_{Ou}$ zur Folge hat.

Ist $\omega \in \mathfrak{P}_{Vx}$, so ist $\pi = \delta^{-1}\omega\delta \in \mathfrak{P}_{Uy}$, also $\mathfrak{P}_{Uy}$ linear transitiv, und es gilt für alle P auf u: $P\pi = P\delta^{-1}\omega\delta = (O, P)\omega\delta = (O, P\omega)\delta = P\omega$. Da jedes Element $\omega' \in \mathfrak{P}_{Vx}$ mit jedem Element $\pi' \in \mathfrak{P}_{Uy}$ kommutiert (s. Folgerung 3 von Satz 1), folgt $(P\omega)\omega' = (P\pi)\omega' = (P\omega')\pi = (P\omega')\omega$ für $P \mathrel{I} u$, da mit P auch $P\omega'$ auf u liegt. Mithin gilt $\omega\omega' = \omega'\omega$, d.h. $\mathfrak{P}_{Vx}$ ist abelsch und damit nach Lemma 9 auch $\mathfrak{P}_{Ou}$.

Der Beweis der Kommutativität der Homologien $\in \mathfrak{P}_{Ou}$ aus der Existenz einer perspektiven Dualität $\in \mathfrak{D}_{Ou}$ kann unter wesentlich schwächeren Voraussetzungen geführt werden. Wir wollen darauf noch im Anhang näher eingehen, vor allem deshalb, weil sich dabei einige Tatsachen über perspektive Kollineationen und Dualitäten ergeben, die für sich von Interesse sind.

Wir müssen nun umgekehrt zeigen, daß die Kommutativität der Homologien $\in \mathfrak{P}_{Ou}$ bei Voraussetzung von P7 die Existenz einer perspektiven Dualität mit O als Zentrum und u als Achse zur Folge hat. Wir beweisen dazu vorerst folgendes Lemma:

Lemma 10. *Unter Voraussetzung von* P7 *zieht die Kommutativität der Homologien* $\in \mathfrak{P}_{Ou}$ *die einfache Transitivität der Gruppe* $\mathfrak{P}_{UO}$ *nach sich.*

Beweis. Zunächst folgt aus P7 nach Satz 1, Folgerung 3, daß die Gruppe $\mathfrak{P}_{ux}$ einfach transitiv ist. Der innere Automorphismus ψ mit

$$\pi^{\psi} = \omega^{-1}\pi\omega \quad \textit{für } \pi \in \mathfrak{P}_{ux} \textit{ und festes } \omega \in \mathfrak{P}_{Vx}$$

der Gruppe $\mathfrak{P}_{ux}$ führt jede Komponente wieder in eine Komponente über und die Gruppe $\mathfrak{P}_{Vx}$ elementweise in sich über, da nach Lemma 9 $\mathfrak{P}_{Vx}$ abelsch ist, also $\omega^{-1}\omega'\omega = \omega'$ für $\omega' \in \mathfrak{P}_{Vx}$ gilt. Mithin sind die nach Satz 1 zu ψ gehörenden Kollineationen λ nach Satz 1, Folgerung 2 Elemente von $\mathfrak{P}_{Uy}$, wenn $P = P'$ und $P \mathrel{I} y$ gewählt wird. Die Gruppe $\mathfrak{P}_{Uy}$ ist dann linear transitiv, denn seien Q, Q' Punkte $\neq P$, U auf (P, U), so ist das Element $\in \mathfrak{P}_{ux}$, welches P in Q überführt, die Translation $\tau_{PQ} \in \mathfrak{P}_{Ux}$. Nach Lemma 8 gibt es dann ein $\omega \in \mathfrak{P}_{Vx}$ mit $\omega^{-1}\tau_{PQ}\omega = \tau_{PQ'}$, woraus $Q\lambda = P\tau_{PQ}^{\psi} = P\omega^{-1}\tau_{PQ}\omega = P\tau_{PQ'} = Q'$ folgt. Da daher die Gruppen $\mathfrak{P}_{Uy}$ und $\mathfrak{P}_{Ux}$ linear transitiv sind, muß nach

der dualen Aussage zu Lemma 4 $\mathfrak{P}_{UO}$ einfach transitiv sein. Damit ist Lemma 10 bewiesen.

Auf Grund des Bewiesenen gibt es also zu jedem $\omega \in \mathfrak{P}_{Vx}$ genau ein $\lambda \in \mathfrak{P}_{Uy}$ mit $\omega^{-1}\pi\omega = \lambda^{-1}\pi\lambda$ für alle $\pi \in \mathfrak{P}_{ux}$. Ordnen wir diese Elemente einander zu, so erhalten wir offenbar einen Isomorphismus χ von $\mathfrak{P}_{Vx}$ und $\mathfrak{P}_{Uy}$. Diesen können wir zu einem Isomorphismus der Gruppen $\mathfrak{P}_{ux}$ und $\mathfrak{P}_{UO}$ erweitern, wenn wir bedenken, daß jedes Element $\pi \in \mathfrak{P}_{ux}$ eindeutig in der Form $\omega\tau$ mit $\omega \in \mathfrak{P}_{Vx}$ und $\tau \in \mathfrak{P}_{Ux}$ und jedes Element $\bar{\pi} \in \mathfrak{P}_{UO}$ eindeutig in der Form $\lambda\tau$ mit $\lambda \in \mathfrak{P}_{Uy}$ und $\tau \in \mathfrak{P}_{Ux}$ darzustellen ist (vgl. S. 337). Die Abbildung

$$\pi^{\varrho} = \omega^{\chi}\tau \quad \text{für } \pi = \omega\tau \in \mathfrak{P}_{ux}$$

ist dann nämlich ein Isomorphismus von $\mathfrak{P}_{ux}$ auf $\mathfrak{P}_{UO}$, denn es ist wegen $\omega^{-1}\tau\omega = (\omega^{\chi})^{-1}\tau\,\omega^{\chi}$:

$$\begin{aligned}(\pi_1\pi_2)^{\varrho} &= (\omega_1\tau_1\omega_2\tau_2)^{\varrho} = (\omega_1\omega_2\omega_2^{-1}\tau_1\omega_2\tau_2)^{\varrho} = (\omega_1\omega_2)^{\chi}(\omega_2^{-1}\tau_1\omega_2)\,\tau_2\\ &= \omega_1^{\chi}\omega_2^{\chi}\left((\omega_2^{\chi})^{-1}\tau_1\omega_2^{\chi}\right)\tau_2 = \omega_1^{\chi}\tau_1\omega_2^{\chi}\tau_2 = \pi_1^{\varrho}\pi_2^{\varrho}.\end{aligned}$$

Nach Satz 3 gibt es zu ϱ eine Dualität δ, die perspektive Dualität mit u als Achse und O als Zentrum ist, wenn $P \mathrel{I} y$ und $p \mathrel{I} V$ gewählt wird, denn es gilt $\tau^{\varrho} = \delta^{-1}\tau\delta = \tau$ für $\tau \in \mathfrak{P}_{Ux}$ und damit $U\delta = x$, $x\delta = U$. Außerdem ist $\omega^{\varrho} = \delta^{-1}\omega\delta \in \mathfrak{P}_{Vy}$ für $\omega \in \mathfrak{P}_{Vx}$, also $V\delta = y$. Umgekehrt folgt aus $y = (V, P)$ und $V\delta = y$: $y\delta = (V\delta, P\delta) = y \cdot p = V$. Ist nun Q ein Punkt $\neq U$ auf u und τ die Translation $\in \mathfrak{P}_{Ux}$ mit $V\tau = Q$, so gilt $Q\delta = V\tau\delta = V\delta\tau = y\tau = (O, Q)$. Ist endlich $\bar{y}$ eine beliebige Gerade $\neq x$ durch O, so gibt es eine Translation $\tau \in \mathfrak{P}_{Ux}$ mit $\bar{y} = y\tau$, und es gilt: $\bar{y}\delta = y\tau\delta = y\delta\tau = V\tau = \bar{y} \cdot u$. Mithin ist δ perspektive Dualität mit O als Zentrum und u als Achse und damit insgesamt bewiesen der

Satz 14. *Eine projektive Ebene ist genau dann als projektive Koordinatenebene über einem kommutativen Körper darstellbar, wenn in ihr Axiom* P7 *und* P8 *gelten.*

Wir bemerken nun noch zum Schluß, daß man in Satz 14 statt P8 auch das folgende schwächere Axiom fordern kann:

P8′. *Es gibt wenigstens eine quasiprojektive Dualität* δ'.

Dabei heißt eine Dualität δ' *quasiprojektiv*, wenn die Punkte P einer Geraden g mit $g\delta' \nmid g$ jeweils durch eine Kollineation α auf die Punkte $P\delta' \cdot g$ auf g abgebildet werden können, d.h. wenn die durch δ' auf einer Geraden g induzierte eineindeutige Abbildung der Punkte auf g in eine Kollineation eingebettet werden kann.

Dem Beweis der Äquivalenz der Axiome P8 und P8′ bei Voraussetzung von Axiom P7 schicken wir einige Hilfssätze voraus.

Lemma 11. *Existiert in einer projektiven Ebene* Π, *in der Axiom* P7 *erfüllt ist, eine Dualität* δ, *so ist in* Π *die Gruppe* $\mathfrak{P}_{Za}$ *bei beliebiger Wahl von* Z *und* a *linear transitiv*[22]).

[22]) Vgl. LENZ [10], Satz 2.

Beweis. (a) *Es gibt eine Kollineation α in Π mit $u\alpha \neq u$ oder $x\alpha \neq x$.* Denn mit $\mathfrak{P}_{xu}$ ist nach Satz 1, Folgerung 3 auch $\mathfrak{P}_{ux}$ einfach transitiv, also sind auch die Gruppen $\delta^{-1}\mathfrak{P}_{xu}\delta = \mathfrak{P}_{u\delta x\delta}$ und $\delta^{-1}\mathfrak{P}_{ux}\delta = \mathfrak{P}_{x\delta u\delta}$ einfach transitiv. Ist nun $u\delta \nmid u$, so kann man für α ein Element $\neq 1$ aus $\mathfrak{P}_{u\delta x\delta}$ mit von u verschiedener Achse wählen und ist $u\delta \nmid x$, so wähle man als Kollineation α eine perspektive Kollineation $\neq 1$ aus $\mathfrak{P}_{u\delta x\delta}$ mit von x verschiedener Achse. Im ersten Fall gilt $u\alpha \neq u$ und im zweiten $x\alpha \neq x$. Ist aber $u\delta = u \cdot x = U$, so ist $x\delta \neq U$, also entweder $x\delta \nmid u$ oder $x\delta \nmid x$, und es gibt stets ein Element $\alpha \neq 1$ aus $\mathfrak{P}_{x\delta u\delta}$ mit $x\alpha \neq x$ oder $u\alpha \neq u$, wie man durch Vertauschen von x und u im vorstehenden Beweis erkennt.

(b) *Es gibt eine Gerade v in Π, so daß die Gruppe $\mathfrak{P}_{Zv}$ für beliebige Wahl von Z linear transitiv ist.* Denn gilt für die nach (a) existierende Kollineation α $x\alpha \neq x$, aber $u\alpha = u$, so folgt (b) für $v = u$ aus Lemma 4, denn mit $\mathfrak{P}_{xu}$ ist auch $\mathfrak{P}_{x\alpha u\alpha} = \mathfrak{P}_{x\alpha u}$ einfach transitiv. Entsprechend folgt für $u\alpha \neq u$, aber $x\alpha = x$ (b) mit $v = x$ aus der Transitivität von $\mathfrak{P}_{ux}$ und $\mathfrak{P}_{u\alpha x}$. Gilt aber $u\alpha \neq u$ und $x\alpha \neq x$ für die nach (a) existierende Kollineation und ist $W = (u\alpha, u)$, so gibt es wegen der Transitivität von $\mathfrak{P}_{u\alpha x\alpha}$ ein Element $\omega \in \mathfrak{P}_{Wx\alpha}$ mit $x\omega \neq x$. Dann ist $\mathfrak{P}_{x\omega u\omega} = \mathfrak{P}_{x\omega u}$ einfach transitiv und damit nach Lemma 4 die Behauptung von (b) für $v = u$ richtig.

Nach [*13*], Abschn. 3, Satz 20 und 27 folgt dann aus (b) sofort Lemma 11.

Lemma 12. *Ist δ eine Dualität, so gibt es eine Gerade p mit $p\delta \nmid p$, wobei p so gewählt werden kann, daß $p\delta \nmid g$ für eine beliebig vorgegebene Gerade g gilt.*

Beweis. Gilt $g\delta \nmid g$, so wähle man für p die Gerade g selbst. Sei also $g\delta \mid g$. Dann gibt es eine Gerade h durch $g\delta^{-1}$ auf g $(h \neq g)$ mit $h\delta \neq g\delta^{-1}$, denn es gibt wenigstens zwei verschiedene Geraden $h, l \neq g$ durch $g\delta^{-1}$, so daß wenigstens eine der beiden verschiedenen Punkte $h\delta$, $l\delta$ ungleich $g\delta^{-1}$ sein muß. Sei ferner p eine Gerade $\neq g$ durch $h\delta$, dann ist $p \neq h$ und $p \cdot h \nmid g$. Seien endlich q, r zwei verschiedene Geraden $\neq p, h$ durch $p \cdot h$, dann kann nicht $p\delta \mid p$ und $q\delta \mid q$ und $r\delta \mid r$ gleichzeitig gelten, da $h\delta, p\delta, q\delta, r\delta$ kollinear und voneinander verschieden sind. Sei etwa $p\delta \nmid p$, dann erfüllt p die Bedingungen von Lemma 12. Ist aber $q\delta \nmid q$ oder $r\delta \nmid r$, so erfüllt q oder r die Bedingungen von Lemma 12.

Lemma 13. *Unter Voraussetzung von Axiom P7 gilt: Ist δ eine quasiprojektive Dualität und sind g, h zwei beliebige Geraden mit $h \nmid g\delta$, so gibt es stets eine Kollineation β mit $P\beta = P\delta \cdot h$ für alle P auf g*[23]).

Beweis. (a) Sei zunächst $g\delta \nmid g$. Dann existiert nach Voraussetzung eine Kollineation α mit $P\alpha = P\delta \cdot g$ für alle P auf g. Nach Lemma 11 läßt sich weiter eine perspektive Kollineation ω mit $g\delta$ als Zentrum und einer mit g, h im Büschel liegenden Achse derart finden, daß g in h übergeführt wird. Dann gilt $P\alpha\omega = P\delta \cdot h$ für alle P auf g, woraus die Behauptung von Lemma 13 mit $\beta = \alpha\omega$ folgt.

[23]) Lemma 13 zeigt, daß in einer projektiven Ebene, in der Axiom P7 gilt, die hier gegebene Definition einer quasiprojektiven Dualität mit der von Lenz in [*11*] angegebenen (spezielleren) Definition übereinstimmt.

(b) Sei $g\delta \mathrel{I} g$, dann gibt es nach Lemma 12 eine Gerade p mit $p\delta \nmid p, g$ und nach (a) eine Kollineation γ mit $Q\gamma = Q\delta \cdot g$ für alle Q auf p. Ist dann P ein Punkt auf g und $P\gamma^{-1} = Q$, so gilt ersichtlich $(P, p\delta) = Q\delta$, woraus $P\delta \cdot p\delta^2 = Q\delta^2 = P\gamma^{-1}\delta^2$ folgt. Es gilt dann $p\delta^2 = q \nmid g\delta$, da sonst $p\delta \mathrel{I} g$ wäre. Sei nun ω die nach Lemma 11 existierende perspektive Kollineation mit $g\delta$ als Zentrum und einer mit h, q im Büschel liegenden Achse, die q in h überführt, dann gilt $P\gamma^{-1}\delta^2\omega = P\delta \cdot h$, also mit $\beta = \gamma^{-1}\delta^2\omega$ die Behauptung von Lemma 13.

Lemma 14. *Unter Voraussetzung von* P7 *gilt: Ist δ eine quasiprojektive Dualität und α eine Kollineation, so ist $\delta\alpha$ bzw. $\alpha\delta$ eine quasiprojektive Dualität.*

Beweis. Wir zeigen zunächst, daß $\delta\alpha$ eine quasiprojektive Dualität ist. Hierfür muß man sich überzeugen, ob für $g\delta\alpha \nmid g$ eine Kollineation β mit $P\beta = P\delta\alpha \cdot g$ existiert. Wegen $g\delta\alpha \nmid g$ ist $g\alpha^{-1} \nmid g\delta$, mithin gibt es nach Lemma 13 eine Kollineation γ mit $P\gamma = P\delta \cdot g\alpha^{-1}$ für alle P auf g, woraus sofort $P\gamma\alpha = P\delta\alpha \cdot g$, also für $\beta = \gamma\alpha$ die Behauptung folgt.

Für $\alpha\delta$ schließen wir entsprechend: Sei $g\alpha = h$, so gibt es wegen $h\delta \nmid g$ nach Lemma 13 eine Kollineation γ mit $Q\gamma = Q\delta \cdot g$ für alle Q auf h, mithin folgt $P\alpha\gamma = P\alpha\delta \cdot g$ für alle P auf g, woraus sich mit $\beta = \alpha\gamma$ die Behauptung ergibt.

Lemma 15. *Unter Voraussetzung von Axiom* P7 *und Axiom* P8′ *gibt es eine quasiprojektive Dualität δ' mit $O\delta' = u$ und $u\delta' = O$.*

Beweis. Nach Lemma 12 gibt es für die nach P8′ existierende quasiprojektive Dualität δ eine Gerade g mit $g\delta \nmid g$. Wir dürfen dann annehmen, daß $g = u$ und $g\delta = O$ ist, denn ist $g \neq u$ und P bzw. Q ein Punkt $\neq g \cdot u$ auf g bzw. u und $(P, g\delta) \cdot (Q, O) = R$, so gibt es nach Lemma 11 eine perspektive Kollineation ω mit einem mit $O, g\delta$ und mit Q, P kollinearen Punkt als Zentrum und einer durch $g \cdot u$ und R gehenden Achse, die $g\delta$ in O und g in u überführt. Dann aber ist $\omega^{-1}\delta\omega$ nach Lemma 14 eine quasiprojektive Dualität δ'' mit $u\delta'' = O$. Es gibt dann weiter eine perspektive Kollineation ω' mit O als Zentrum, die $u\delta''^2$ in u überführt. Dann ist $\delta''\omega'$ nach Lemma 14 eine quasiprojektive Dualität δ' mit $u\delta' = u(\delta''\omega') = O\omega' = O$ und $O\delta' = (u\delta'')\,\delta''\omega' = u\delta''^2\omega' = u$, wie verlangt.

Beweis der Äquivalenz von P8 und P8′ bei Voraussetzung von P7:

(a) *Aus* P8′ *folgt* P8: Sei δ' die nach Lemma 15 vorhandene quasiprojektive Dualität und α eine Kollineation mit $P\alpha = P' = P\delta' \cdot u$ für alle P auf u. Wir dürfen annehmen, daß $O\alpha = O$ ist [anderenfalls gibt es nach P7 und Lemma 4 eine perspektive Kollineation π mit u als Achse, für die $(O\alpha)\pi = O$ ist, und man kann α durch $\alpha' = \alpha\pi$ ersetzen, da für α' dann ebenfalls $P\alpha' = P\alpha\pi = P'\pi = P'$ für $P \mathrel{I} u$, aber außerdem $O\alpha' = O$ gilt]. $\delta'\alpha^{-1}$ ist dann perspektive Dualität $\in \mathfrak{D}_{Ou}$, denn es gilt $P\delta'\alpha^{-1} = (P', O)\alpha^{-1} = (P, O)$ für alle $P \mathrel{I} u$ und $g\delta'\alpha^{-1} = (g, u) = Q$ für $g \mathrel{I} O$; letzteres, da sich aus $g = (O, Q)$ ergibt $g\delta' = u \cdot Q\delta' = u \cdot (Q', O)$.

(b) *Aus* P8 *folgt* P8′: Sei δ die nach P8 existierende Dualität $\in \mathfrak{D}_{Ou}$, so ist δ quasiprojektiv, da für die Gerade u $\alpha = 1$ gewählt werden kann und

für $g \neq u$ $[g \nmid O]$ die Translation $\alpha = \tau_{OG} \in \mathfrak{P}_{uu}$ mit $G = g\delta$ das Verlangte leistet. Die Translation τ_{OG} existiert, da nach Satz 14 und 10 die Gruppe $\mathfrak{P}_{uu}$ einfach transitiv ist. Wir haben damit bewiesen den

Satz 15. *Eine projektive Ebene ist genau dann als projektive Koordinatenebene über einem kommutativen Körper darstellbar, wenn in ihr Axiom* P7 *und* P8′ *erfüllt sind.*

Satz 15 entnimmt man, daß in der projektiven Koordinatenebene über einem Fastkörper $\mathfrak{F}$ genau dann eine quasiprojektive Dualität existiert, wenn $\mathfrak{F}$ ein kommutativer Körper ist.

Anhang

Wir wollen nun den in § 5 angekündigten Beweis darstellen. Zunächst leiten wir her den

Satz 16. *Gelten in einer projektiven Ebene die Axiome* P7b *und* P8, *so gibt es eine Homologie* $\sigma \in \mathfrak{P}_{Ou}$ *mit*

$$\delta^2 = \sigma \quad \text{für alle } \delta \in \mathfrak{D}_{Ou}. \tag{31}$$

σ ist entweder involutorische Homologie $\in \mathfrak{P}_{Ou}$ *oder die Identität, und es gilt* $\sigma\tau\sigma = \tau^{-1}$ *für alle* $\tau \in \mathfrak{P}_{Uu}$.

Beweis. Sei δ ein beliebiges Element $\in \mathfrak{D}_{Ou}$ und τ eine beliebige Translation $\neq 1$ aus $\mathfrak{P}_{Uu}$ und $O\tau = A$. Dann ist $\delta^{-1}\tau^{-1}\delta = \tau'$ eine Translation $\in \mathfrak{P}_{Ox}$ mit $x = (O, U)$, für die $(A\delta)\tau' = A\,\delta\delta^{-1}\tau^{-1}\delta = u$ gilt. Sei dann weiter B ein beliebiger Punkt $\neq U$ auf $A\delta$, dann muß, da $B\tau'$ auf u liegt,

$$B\delta^{-1} = B\tau'\delta^{-1}\tau = (O, B\tau')\tau = (A, B\tau) \tag{32}$$

und wegen $(O, B\tau') = (O, B)$ auch

$$B\delta^{-1} = (O, B)\tau = (A, B\tau) \tag{33}$$

gelten. Mit B liegt nun auch $B\tau^{-1}$ auf $A\delta$, da $A\delta$ mit U inzidiert. Wendet man daher auf $B\tau^{-1}$ die Gl. (33) an, so hat man $(B\tau^{-1})\delta^{-1} = (A, B)$. Andererseits ist nach (32) auch $B\delta^{-1}\tau'^{-1} = (A, B)$, so daß $B\tau^{-1}\delta^{-1} = B\delta^{-1}\tau'^{-1}$ oder $B\delta\tau^{-1}\delta^{-1} = B\tau'^{-1}$ folgt. Da $\delta\tau^{-1}\delta^{-1} \in \mathfrak{P}_{Ox}$ ist, muß nach Lemma 1 $\delta\tau^{-1}\delta^{-1} = \tau'^{-1}$, also $\delta\tau^{-1}\delta^{-1} = \delta^{-1}\tau\delta$ und endlich

$$(\delta^2)^{-1}\tau\delta^2 = \tau^{-1} \tag{34}$$

gelten. Da diese Gleichung für beliebiges $\delta \in \mathfrak{D}_{Ou}$ richtig ist, muß δ^2 für alle $\delta \in \mathfrak{D}_{Ou}$ gleich einer festen Homologie $\sigma \in \mathfrak{P}_{Ou}$, mithin (31) gültig sein, denn wegen der Eindeutigkeitsaussage in Lemma 8 (die nur Lemma 1, nicht P7 verwendet) gibt es höchstens eine Homologie $\in \mathfrak{P}_{Ou}$ mit $\sigma^{-1}\tau\sigma = \tau^{-1}$.

Aus $\sigma^{-1}\tau\sigma = \tau^{-1}$ folgt $\sigma^{-2}\tau\sigma^2 = \tau$ und daher nach der Eindeutigkeitsaussage in Lemma 8 $\sigma^2 = 1$. Also ist σ entweder involutorisch oder die Identität. Aus (34) entsteht dann die Gleichung

$$\delta^2\tau\delta^2 = \sigma\tau\sigma = \tau^{-1} \quad \text{für alle } \delta \in \mathfrak{D}_{Ou} \text{ und alle } \tau \in \mathfrak{P}_{Uu}, \tag{35}$$

womit Satz 16 vollständig bewiesen ist.

(31) entnimmt man im übrigen, daß alle Elemente $\in \mathfrak{D}_{O_u}$ die gleiche Ordnung, und zwar höchstens die Ordnung 4 haben. Mit Satz 16 erhält man ohne Mühe den

Satz 17. *Unter Voraussetzung von Axiom* P7b *und* P8 *gilt*

$$\delta^{-1}\omega\delta = \omega^{-1} \quad \textit{für alle } \delta \in \mathfrak{D}_{O_u} \textit{ und alle } \omega \in \mathfrak{P}_{O_u}. \tag{36}$$

Beweis. Mit δ ist offensichtlich auch $\delta\omega \in \mathfrak{D}_{O_u}$ und daher nach Satz 16 $\delta^2 = \sigma$ und $(\delta\omega)^2 = \sigma$, also $\delta^{-1} = \sigma\delta$ und $\sigma\delta\omega\delta\omega = 1$, was zusammen $\delta^{-1}\omega\delta = \omega^{-1}$ ergibt.

Aus (35) und (36) folgt dann, daß $\mathfrak{P}_{U_u}$ und $\mathfrak{P}_{O_u}$ abelsch sind, denn sind τ_1, τ_2 zwei Translationen $\in \mathfrak{P}_{U_u}$, so gilt wegen (35): $\tau_1\tau_2 = \sigma(\tau_1\tau_2)^{-1}\sigma = \sigma\tau_2^{-1}\tau_1^{-1}\sigma = \sigma\tau_2^{-1}\sigma\sigma\tau_1^{-1}\sigma = \tau_2\tau_1$. Entsprechend schließt man aus (36), daß die Homologien $\in \mathfrak{P}_{O_u}$ kommutieren. Damit können wir abschließend formulieren:

Satz 18. *Sind in einer projektiven Ebene die Axiome* P7b *und* P8 *erfüllt, so sind die Gruppen* $\mathfrak{P}_{U_u}$ *und* $\mathfrak{P}_{O_u}$ *abelsch.*

Literatur

[1] ANDRÉ, J.: Über nicht-Desarguessche Ebenen mit transitiver Translationsgruppe. Math. Z. **60**, 156—186 (1954). — [2] ANDRÉ, J.: Projektive Ebenen über Fastkörpern. Math. Z. **62**, 137—160 (1955). — [3] ARTIN, E.: Geometric Algebra. Interscience tracts in pure and applied mathematics **3**. New York 1957. — [4] BAER, R.: Homogeneity of projective planes. Amer. J. Math. **64**, 137—152 (1942). — [5] BAER, R.: Linear algebra and projective geometry. New York 1952. — [6] HALL, M.: Projective planes. Trans. Amer. Math. Soc. **54**, 229—277 (1943). — [7] HALL, M.: Corrections to „Projective planes". Trans. Amer. Math. Soc. **65**, 473—474 (1949). — [8] HESSENBERG, G.: Grundlagen der Geometrie. Berlin 1930. — [9] HILBERT, D.: Grundlagen der Geometrie, 8. Aufl. Stuttgart 1956. — [10] LENZ, H.: Kleiner Desarguesscher Satz und Dualität in projektiven Ebenen. Jber. dtsch. Math. Verein. **57**, 20—31 (1954). — [11] LENZ, H.: Zur Begründung der analytischen Geometrie. Sitzgsber. Bayr. Akad. Wiss., math.-naturw. Kl. **1954**, 17—72. — [12] LINGENBERG, R.: Begründung der absoluten Geometrie der Ebene. Phil. Diss. Kiel 1955. — [13] PICKERT, G.: Projektive Ebenen. Berlin usw. 1955. — [14] SCHWAN, W.: Streckenrechnung und Gruppentheorie. Math. Z. **3**, 11—28 (1919). — [15] VEBLEN, O., and W. YOUNG: Projective geometry, Bd. 1, 2. Boston 1910—1916.

Institut A für Höhere Mathematik an der Technischen Hochschule Hannover

(Eingegangen am 22. November 1956)

WALTER, W.
Math. Zeitschr. Bd. 67, S. 361—376 (1957)

Über die Euler-Poisson-Darboux-Gleichung

Von
WOLFGANG WALTER

1. Einleitung. Die vorliegende Arbeit behandelt das Anfangswertproblem für die Euler-Poisson-Darboux-Gleichung (EPD-Gleichung)

$$L(u) = u_{rr} + \frac{k}{r} u_r - \Delta u = 0. \tag{1.1}$$

Dabei ist $u(r, x_1, \ldots, x_n) = u(r, x)$ eine von $n+1$ reellen Veränderlichen ($n \geq 1$) abhängende reelle Funktion, $\Delta u = u_{x_1 x_1} + \cdots + u_{x_n x_n}$ und k eine reelle Konstante. Unter einer Lösung der Gl. (1.1) verstehen wir immer eine für alle $r > 0$, $x \in E_n$ (E_n ist der n-dimensionale Euklidische Raum) erklärte und zweimal stetig differenzierbare Funktion u, für die $L(u) = 0$ ist. Die Ausdrucksweisen „u besitzt stetige partielle Ableitungen bis zur Ordnung i", „u ist i-mal stetig differenzierbar", „u ist aus der Klasse C^{i}" besagen ein und dasselbe.

Das Cauchysche Anfangswertproblem besteht darin, zu einer vorgegebenen Funktion $f(x)$ (welche gewisse, später noch genau anzugebende Regularitätsvoraussetzungen erfüllen soll) eine Lösung $u(r, x)$ von (1.1) zu finden, für die

$$u(0, x) = f(x) \quad \text{und} \quad u_r(0, x) = 0 \tag{1.2}$$

ist (vgl. jedoch Nr. 5)*).

Es ist für viele Rechnungen zweckmäßig, an Stelle von r die neue unabhängige Veränderliche $s = r^2$ einzuführen. Setzen wir dementsprechend $\bar{u}(s, x) = u(\sqrt{s}, x)$, so wird aus (1.1) die Gleichung

$$4 s \bar{u}_{ss} + 2(k+1) \bar{u}_s - \Delta \bar{u} = 0$$

oder

$$\bar{L}(\bar{u}) = s \bar{u}_{ss} + c \bar{u}_s - \frac{1}{4} \Delta \bar{u} = 0, \qquad c = \frac{k+1}{2}; \tag{1.3}$$

die Randbedingungen lauten jetzt

$$\bar{u}(0, x) = f(x) \quad \text{und} \quad \sqrt{s}\, \bar{u}_s(s, x) \to 0 \quad \text{für } s \to 0. \tag{1.4}$$

Die letzte Bedingung ist insbesondere dann erfüllt, wenn $\bar{u}_s$ für kleine s beschränkt bleibt.

Eine vollständige Lösung des Anfangswertproblems gaben WEINSTEIN [9] und später mit anderer Methode DIAZ und WEINBERGER [4]. Beide Autoren

*) Genauer: u und u_r sollen für $r \geq 0$ stetig sein und den Gln. (1.2) genügen. Gleichbedeutend sind die Formulierungen: Für $(r, x) \to (0, x_0)$ soll $u(r, x) \to f(x_0)$, $u_r(r, x) \to 0$ streben; für $r \to 0$ soll — gleichmäßig in jedem beschränkten x-Bereich — $u(r, x) \to f(x)$, $u_r(r, x) \to 0$ streben.

gehen vom Fall $k \geq n-1$ aus, wo die Lösung durch gewisse Mittelwerte der Funktion $f(x)$ dargestellt werden kann (s. Nr. 3). Durch analytische Fortsetzung dieser Formeln gelangen DIAZ und WEINBERGER zu Lösungen für $k<n-1$, während bei WEINSTEIN [9], [10], [11] die beiden Rekursionsformeln

$$u_r^{(k)}(r,x) = r\,u^{(k+2)}(r,x); \qquad u^{(k)}(r,x) = r^{1-k}\,u^{(2-k)}(r,x) \tag{1.5}$$

[mit $u^{(k)}$ wird eine Lösung von (1.1) bezeichnet] die Grundlage für die Behandlung der Fälle $k<n-1$ bilden. Die „Ausnahmewerte" $k=-1,-3,-5,\ldots$ $(c=0,-1,-2,\ldots)$ bedürfen einer gesonderten Betrachtung (vgl. dazu auch BLUM [2]). Die Formeln (1.5) sind äquivalent mit

$$\bar{u}_s^{(c)}(s,x) = \bar{u}^{(c+1)}(s,x); \qquad \bar{u}^{(c)}(s,x) = s^{1-c}\,\bar{u}^{(2-c)}(s,x); \tag{1.6}$$

dabei ist $\bar{u}^{(c)}$ eine Lösung von (1.3).

In dieser Arbeit wird zunächst in Nr. 2 das Anfangswertproblem für analytische Anfangswerte behandelt. Die Lösung kann dann als unendliche Reihe dargestellt werden. An diesen Reihen wird unter anderem die schon von WEINSTEIN [9] bemerkte Sonderstellung der polyharmonischen Anfangswerte unmittelbar sichtbar. Bei der Behandlung der Fälle $k \geq n-1$ wurde bisher immer $f \in C^2$ vorausgesetzt; in Nr. 3 wird ein neuer Beweis gegeben, welcher mit geringeren Voraussetzungen auskommt. Genauere Aussagen über das Randverhalten der Lösungen enthält die Nr. 4. Als wichtigstes Ergebnis der letzten Nummer, in der auf Eindeutigkeitsfragen eingegangen wird, sei die Tatsache genannt, daß für $k \neq 0$ die zweite Gl. (1.2) eine Folge der ersten dieser Gleichungen und der EPD-Gl. (1.1) ist, daß also diese zweite Gleichung bei der Stellung des Anfangswertproblems (für $k \neq 0$) entbehrt werden kann.

2. Analytische Anfangswerte. Es werden — nur in dieser Nummer — folgende Bezeichnungen benutzt. Ist $p=(p_1,\ldots,p_n)$ ein Vektor mit ganzzahligen nichtnegativen p_i, so sei $|p| = p_1+\cdots+p_n$ und $x^p = x_1^{p_1}x_2^{p_2}\ldots x_n^{p_n}$. Die Funktion $f(x)=f(x_1,\ldots,x_n)$ wird analytisch genannt, wenn sie eine Darstellung durch eine für alle x absolut konvergente Potenzreihe

$$f(x) = \sum_p a_p\,x^p = \sum_{p_1=0}^{\infty} a_{p_1,\ldots,p_n}\,x_1^{p_1}\ldots x_n^{p_n} \tag{2.1}$$

besitzt. In diesem Falle läßt sich das Anfangswertproblem für die EPD-Gleichung durch Reihenentwicklungen lösen. Die Formeln lassen sich einfacher schreiben, wenn wir die Differentialgleichung (1.3) und dementsprechend Lösungen $\bar{u}(s,x)$ betrachten.

Satz 2.1. *Ist die Funktion $f(x)$ analytisch und ist $c \neq 0,-1,-2,\ldots$, so wird durch die Reihe*

$$\bar{u}(s,x) = \sum_{\nu=0}^{\infty} \frac{s^{\nu}\,\Delta^{\nu} f(x)}{4^{\nu}\,\nu!\,(c)_{\nu}} \tag{2.2}$$

mit

$$(c)_0 = 1, \quad (c)_{\nu} = c(c+1)\ldots(c+\nu-1) \quad \text{für } \nu \geq 1 \tag{2.3}$$

eine Lösung der Gl. (1.3) *bestimmt, welche für* $s=0$ *die Werte*

$$\bar{u}(0,x)=f(x) \tag{2.4}$$

$$\frac{\partial^i \bar{u}(0,x)}{\partial s^i}=\frac{\Delta^i f(x)}{4^i (c)_i} \qquad (i=1,2,\ldots) \tag{2.5}$$

annimmt. Die Reihe (2.2) *ist eine für alle* s, x *absolut konvergente Potenzreihe. Ihre Summe stellt also (da* $\bar{u}_s$ *für kleine* s *beschränkt ist) eine Lösung des Anfangswertproblems* (1.3), (1.4) *dar.*

Beweis. Die Reihe $\sum_{i=0}^{\infty} A_i t^i$ mit

$$A_i=\sum_{|p|=i}|a_p| \qquad (i=0,1,2,\ldots)$$

hat nach unseren Voraussetzungen den Konvergenzradius ∞. Setzen wir

$$\Delta f(x)=\sum_p a_p^1 x^p \quad \text{und} \quad A_i^1=\sum_{|p|=i}|a_i^1|,$$

so ergibt wegen

$$\Delta x^p=p_1(p_1-1)\,x_1^{p_1-2}x_2^{p_2}\ldots x_n^{p_n}+\cdots+p_n(p_n-1)\,x_1^{p_1}\ldots x_{n-1}^{p_{n-1}}x_n^{p_n-2}$$

eine einfache Abschätzung, daß

$$A_i^1\leq n(i+1)(i+2)\,A_{i+2}$$

ist. Daraus folgt sofort, daß, wenn

$$\Delta^\nu f(x)=\sum_p a_p^\nu x^p \quad \text{und} \quad A_i^\nu=\sum_{|p|=i}|a_p^\nu|$$

gesetzt wird, die Ungleichung

$$A_i^\nu\leq n^\nu(i+1)(i+2)\ldots(i+2\nu)\,A_{i+2\nu} \tag{2.6}$$

besteht.

Die Reihe (2.2) läßt also für $|x_i|\leq t$, $|s|\leq t^2$ die Abschätzung

$$|\bar{u}(s,x)|\leq\sum_{\nu=0}^{\infty}\frac{t^{2\nu}}{4^\nu\nu!\,|(c)_\nu|}\sum_p|a_p^\nu|\,t^{|p|}=\sum_{\nu=0}^{\infty}\frac{t^{2\nu}}{4^\nu\nu!\,|(c)_\nu|}\sum_{i=0}^{\infty}A_i^\nu t^i$$

$$\leq\sum_{\nu=0}^{\infty}\frac{t^{2\nu}}{4^\nu\nu!\,|(c)_\nu|}\sum_{i=0}^{\infty}n^\nu(i+1)(i+2)\ldots(i+2\nu)\,A_{i+2\nu}t^i$$

$$\leq\sum_{j=0}^{\infty}A_j t^j\sum_{\nu=0}^{\infty}\frac{n^\nu j(j-1)\ldots(j-2\nu+1)}{4^\nu\nu!\,|(c)_\nu|}$$

zu. Man kann, da $c\neq 0,-1,\ldots$ sein sollte, immer eine positive Konstante α so wählen, daß für alle $\nu\geq 0$

$$\alpha^\nu|(c)_\nu|\geq\nu!, \quad \text{also} \quad \frac{1}{4^\nu\nu!\,|(c)_\nu|}\leq\frac{\alpha^\nu}{(2\nu)!}$$

ist. Dann ist aber

$$\sum_{\nu} \frac{n^{\nu} j(j-1)\dots(j-2\nu+1)}{4^{\nu}\nu!\,|(c)_{\nu}|} \leq \sum_{\nu} (n\alpha)^{\nu} \binom{j}{2\nu} \leq (n\alpha)^{j} \sum_{\mu=0}^{j} \binom{j}{\mu} = (2n\alpha)^{j},$$

also

$$|\bar{u}(s,x)| \leq \sum_{i=0}^{\infty} A_i (2n\alpha)^i t^i,$$

d.h. die Reihe (2.2) ist für alle s, x absolut konvergent.

Der Rest des Beweises ist einfach. Durch gliedweise Differentiation ergibt sich

$$\begin{aligned} \bar{L}(\bar{u}) &= \sum_{\nu=0}^{\infty} \frac{s^{\nu-1}\Delta^{\nu} f}{4^{\nu}\nu!\,(c)_{\nu}} [\nu(\nu-1) + c\nu] - \sum_{\nu=0}^{\infty} \frac{s^{\nu}\Delta^{\nu+1} f}{4^{\nu+1}\nu!\,(c)_{\nu}} \\ &= \sum_{\nu=1}^{\infty} \frac{s^{\nu-1}\Delta^{\nu} f}{4^{\nu}\nu!\,(c)_{\nu}} [\nu(\nu+c-1) - \nu(\nu+c-1)] = 0. \end{aligned}$$

Schließlich erhält man die Formeln (2.4), (2.5) sofort aus (2.2).

Ist c eine der Zahlen $0, -1, -2, \dots$ (d.h. k eine der Zahlen $-1, -3, \dots$), so verliert die Reihe (2.2) ihren Sinn, da im Nenner Nullen auftreten. Doch lassen sich auch für diese Fälle Reihenentwicklungen angeben.

Satz 2.2. *Ist $c = -q$ eine ganze Zahl ≤ 0, so ist die Funktion*

$$\bar{u}(s,x) = \sum_{\nu=0}^{q} \frac{s^{\nu}\Delta^{\nu} f(x)}{4^{\nu}\nu!\,(c)_{\nu}} + \sum_{\nu=q+1}^{\infty} \frac{s^{\nu}\Delta^{\nu} f(x)\,(\log s - \beta_{\nu})}{4^{\nu}\nu!\,(c)_{\nu}^{*}} \tag{2.7}$$

eine Lösung der EPD-Gl. (1.3). *Dabei soll*

$$\beta_{q+1} = 0, \quad \beta_{q+i} = \sum_{j=1}^{i-1} \left(\frac{1}{j} + \frac{1}{q+j+1} \right) \quad (i = 2, 3, \dots) \tag{2.8}$$

und $(c)_{\nu}^{*}$ die Zahl sein, welche sich ergibt, wenn man den im Produkt $(c)_{\nu}$ auftretenden Faktor 0 wegläßt, also

$$(-q)_{\nu}^{*} = (-1)^{q} q!\,(\nu-q-1)! \quad \text{für } \nu \geq q+1. \tag{2.9}$$

Die Funktion $\bar{u}$ nimmt die Randwerte

$$\frac{\partial^{i}\bar{u}(0,x)}{\partial s^{i}} = \frac{\Delta^{i} f(x)}{4^{i}(c)_{i}} \quad \text{für } i = 0, 1, \dots, q \tag{2.10}$$

an, während die $(q+1)$-te Ableitung (neben anderen, für $s \to 0$ beschränkt bleibenden Gliedern) ein Glied mit einer logarithmischen Singularität

$$\frac{(-1)^{q}\Delta^{q+1} f(x) \log s}{4^{q+1} q!} \tag{2.11}$$

aufweist. Die (für alle x und $s > 0$ absolut konvergente) Reihe (2.7) *ist also eine Lösung des Anfangswertproblems* (1.3), (1.4).

Beweis. Der Konvergenzbeweis kann, da für $\nu \geq q+1$

$$\frac{1}{4^\nu \nu!\,|(c)_\nu^*|} = \frac{1}{4^\nu \nu!\,q!\,(\nu-q-1)!} \leq \frac{\nu^{q+1}}{(2\nu)!} \leq \frac{\alpha^\nu}{(2\nu)!}$$

bei geeignetem $\alpha > 0$ und da $\beta_\nu < 2^\nu$ ist, ganz ähnlich wie oben bei Satz 1 durchgeführt werden.

Der Koeffizient von $s^{\nu-1}\Delta^\nu \bar{u}$ lautet in der Reihe für $\Delta\bar{u}/4$

$$(2.12)\qquad \frac{1}{4^\nu(\nu-1)!\,(c)_{\nu-1}} \quad \text{für } \nu \leq q+1; \qquad \frac{\log s - \beta_{\nu-1}}{4^\nu(\nu-1)!\,(c)^*_{\nu-1}} \quad \text{für } \nu \geq q+2,$$

und in der Reihe für $s\bar{u}_{ss} + c\bar{u}_s$

$$(2.13)\qquad \begin{cases} \dfrac{\nu(\nu-1)+c\nu}{4^\nu \nu!\,(c)_\nu} & \text{für } \nu \leq q; \\ \dfrac{\nu(\nu+c-1)(\log s - \beta_\nu) + 2\nu + c - 1}{4^\nu \nu!\,(c)^*_\nu} & \text{für } \nu \geq q+1. \end{cases}$$

Die Koeffizienten aus (2.12) und (2.13) stimmen, wie man sofort sieht, für $\nu \leq q$ überein. Für $\nu = q+1$ ergibt der erste Ausdruck in (2.12) und der zweite Ausdruck in (2.13) denselben Wert $[4^{q+1}q!\,(c)_q]^{-1}$, für $\nu \geq q+2$ muß die Gültigkeit von

$$\frac{1}{4^\nu(\nu-1)!\,(c)^*_{\nu-1}} = \frac{\nu(\nu+c-1)}{4^\nu \nu!\,(c)^*_\nu} \quad \text{und} \quad \frac{-\beta_{\nu-1}}{4^\nu(\nu-1)!\,(c)^*_{\nu-1}} = \frac{-\nu(\nu+c-1)\beta_\nu + 2\nu + c - 1}{4^\nu \nu!\,(c)^*_\nu}$$

nachgewiesen werden. Hiervon ist die erste Gleichung offenbar richtig, während die zweite mit

$$\beta_\nu = \beta_{\nu-1} + \frac{2\nu+c-1}{\nu(\nu+c-1)} = \beta_{\nu-1} + \frac{1}{\nu} + \frac{1}{\nu+c-1}$$

gleichwertig, also nach (2.8) ebenfalls richtig ist.

Die Behauptungen über die Randwertannahme folgen sofort aus (2.7).

Die Lösung des Anfangswertproblems ist für $c \geq \frac{1}{2}$ $(k \geq 0)$ eindeutig bestimmt ([3, S. 381]; vgl. Nr. 5), während für $c < \frac{1}{2}$ nach der zweiten Formel in (1.6) $s^{1-c}\,\bar{u}^{(2-c)}(s, x)$, insbesondere also

$$s^{1-c}\sum_{\nu=0}^{\infty} \frac{s^\nu \Delta^\nu g(x)}{4^\nu \nu!\,(2-c)_\nu}$$

bei beliebigem analytischem $g(x)$ eine Lösung $\bar{u}^{(c)}$ ist, für die die Anfangswerte (1.4) verschwinden.

Für die Wellengleichung $\Delta u(r, x) = u_{rr}(r, x)$ lautet (2.2)

$$u(r, x) = \sum_{\nu=0}^{\infty} \frac{r^{2\nu}\Delta^\nu f(x)}{(2\nu)!}.$$

Die Lösungen (2.2) und (2.7) werden besonders einfach, wenn $f(x)$ eine polyharmonische Funktion ist. Aus $\Delta^i f(x) = 0$ folgt nämlich, daß die Reihen bei $\nu = i$ abbrechen. Liegt der Fall von Satz 2.2 vor und ist dabei noch

$i \leq q+1$, so verschwinden die logarithmischen Glieder und damit auch die logarithmische Singularität der Lösung. Diese zuletzt genannte Tatsache wurde auf anderem Wege von WEINSTEIN bewiesen ([9], [10, S. 146]; vgl. auch [4, S. 711]).

3. Der Fall $k \geq n-1$ ($c \geq n/2$). Es ist bekannt, daß in diesen Fällen die Lösung des Anfangswertproblems explizit als ein geeigneter Mittelwert der Funktion $f(x)$ angegeben werden kann ([4, S. 705–707], [9, S. 143]). Wir geben hier einen neuen Beweis mit abgeschwächten Voraussetzungen bezüglich $f(x)$ [1]).

Zunächst einige Erklärungen. Wir nennen $f(x)$ lokalsummierbar (kurz $f \in L_0$), wenn f über jedes beschränkte Gebiet des E_n integrierbar ist. Der Oberflächenmittelwert von $f(x)$ (bezüglich einer Kugel vom Mittelpunkt x und Radius r) ist erklärt durch

$$M(r,x,f) = \frac{1}{\omega r^{n-1}} \int\limits_{|y-x|=r} f(y)\,dy = \frac{1}{\omega} \int\limits_{|y|=1} f(x+ry)\,dy; \tag{3.1}$$

dabei ist $\omega = 2\pi^m/\Gamma(m)$, $m = n/2$, der Oberflächeninhalt der Einheitskugel und dy jeweils das Oberflächenelement auf der unter dem Integralzeichen angegebenen Fläche (wir benutzen dasselbe Zeichen für Raumelement und Oberflächenelement; dy bedeutet das eine oder das andere, je nachdem, ob das in Rede stehende Integral ein Raum- oder ein Oberflächenintegral ist). Ist $f \in L_0$, so existiert $M(r,x,f)$ bei festem x für fast alle $r>0$ (bei stetigem f natürlich für alle $r>0$), und $r^{n-1}M(r,x,f)$ ist über jedes (endliche) r-Intervall integrierbar [2]).

Daneben werden von einem Parameter $q>0$ abhängende Mittelwerte betrachtet; ihre Definition lautet (in mehreren gleichbedeutenden Formen)

$$\left\{\begin{aligned} M(r,x,f,q) &= \frac{2}{\omega\, B(m,q)\, r^n} \int\limits_{|y-x|<r} f(y)\left(1-\frac{|y-x|^2}{r^2}\right)^{q-1} dy \\ &= \frac{2}{\omega\, B(m,q)} \int\limits_{|y|<1} f(x+ry)\,(1-|y|^2)^{q-1}\,dy \\ &= \frac{2}{B(m,q)\, r^n} \int\limits_0^r M(\varrho,x,f)\,\varrho^{n-1}\left(1-\frac{\varrho^2}{r^2}\right)^{q-1} d\varrho, \end{aligned}\right. \tag{3.2}$$

wobei

$$m = n/2 \quad \text{und} \quad B(m,q) = \frac{\Gamma(m)\,\Gamma(q)}{\Gamma(m+q)} = \int\limits_0^1 t^{m-1}(1-t)^{q-1}\,dt \tag{3.3}$$

das Eulersche Integral 1. Gattung ist. Bei stetigem $f(x)$ existiert dieser Mittelwert für $q>0$, für $f \in L_0$ sicher für $q \geq 1$.

[1]) In den uns zugänglichen Arbeiten wird $f \in C^2$ vorausgesetzt; lediglich GHERMANESCU [5, S. 291 f.] beweist (1.1) für $k = n+1, n+3, n+5, \ldots$ und $f \in C^1$.

[2]) Vgl. WALTER [8, S. 96].

Auch hier ist es manchmal zweckmäßig, statt der Mittelwerte $M(r, \ldots)$ die Funktionen $\overline{M}(s, \ldots) = M(\sqrt{s}, \ldots)$ zu betrachten. Für die letzteren gilt nach (3.1) und (3.2)

$$\overline{M}(s, x, f) = \frac{1}{\omega} \int_{|y|=1} f(x + \sqrt{s}\, y)\, d\, y \tag{3.4}$$

und

$$\overline{M}(s, x, f, q) = \frac{1}{B(m, q)\, s^{m+q-1}} \int_0^s \overline{M}(\sigma, x, f)\, \sigma^{m-1} (s - \sigma)^{q-1} d\sigma. \tag{3.5}$$

Der bei den Mittelwerten vor dem Integralzeichen stehende Faktor ist so gewählt worden, daß — wie man aus einem Vergleich von (3.3) und (3.5) erkennt — für $f(x) = 1$ auch $M(r, \ldots) = 1$ für alle betrachteten Mittelwerte ist. Daraus folgt weiter

$$M(r, x, f) \to f(x) \quad \text{und} \quad M(r, x, f, q) \to f(x) \tag{3.6}$$

für stetiges f (dasselbe gilt natürlich für $\overline{M}$ bei $s \to 0$). Um im folgenden Fallunterscheidungen zu vermeiden, führen wir die Definition

$$M(r, x, f) = M(r, x, f, 0); \qquad \overline{M}(s, x, f) = \overline{M}(s, x, f, 0) \tag{3.7}$$

ein. Sie findet ihre Berechtigung in der (leicht zu beweisenden, aber im folgenden nicht benötigten) Tatsache, daß $M(r, x, f, q) \to M(r, x, f)$ strebt für $q \to 0$, falls f stetig ist.

Satz 3.1. *Ist $f(x)$ stetig und i-mal stetig differenzierbar ($i = 0, 1, \ldots$), so gilt dasselbe für $M(r, x, f, q)$, $q \geq 0$, als Funktion von x (bei festem r), und es ist, wenn $D^j = \partial^j / \partial x_1^{j_1} \ldots \partial x_n^{j_n}$, $j = j_1 + \cdots + j_n$, irgendeine partielle Ableitung der Ordnung $j \leq i$ bedeutet,*

$$D^j M(r, x, f, q) = M(r, x, D^j f, q). \tag{3.8}$$

In der zweiten Integraldarstellung (3.1) bzw. (3.2) darf man nämlich unter dem Integralzeichen differenzieren. Schärfere Ergebnisse werden wir aus dem folgenden Hilfssatz gewinnen.

Hilfssatz 3.1. *Die Funktion $F(t, y)$ sei, wenn die reelle Veränderliche t das Intervall $I: t_0 \leq t \leq t_1$ und der Vektor y den E_n durchläuft, meßbar und für jedes feste y als Funktion von t in I totalstetig. Ferner mögen die L-Integrale*

$$G(t) = {}^{E_n}\!\int F(t, y)\, d\, y, \quad H(t) = {}^{E_n}\!\int F_t(t, y)\, d\, y, \quad J(t) = {}^{E_n}\!\int |F_t(t, y)|\, d\, y \tag{3.9}$$

für jedes $t \in I$ existieren und die Funktionen $H(t)$, $J(t)$ in I stetig sein. Dann ist $G(t)$ in I stetig differenzierbar und

$$G_t(t) = {}^{E_n}\!\int F_t(t, y)\, d\, y. \tag{3.10}$$

Beweis. Nach den Sätzen von Fubini und Tonelli [6, S. 151] ist

$$\int_{t_0}^{t} H(z)\, dz = \int_{t_0}^{t} {}^{E_n}\!\int F_t(z, y)\, d\, y\, dz = {}^{E_n}\!\int \int_{t_0}^{t} F_t(z, y)\, dz\, d\, y$$

$$= {}^{E_n}\!\int [F(t, y) - F(t_0, y)]\, d\, y = G(t) - G(t_0),$$

also $H(t) = G_t(t)$.

Wird unter $(1-z)_*^q$ die Funktion

$$(3.11) \qquad (1-z)_*^q = \begin{cases} (1-z)^q & \text{für } 0 \leq z \leq 1 \\ 0 & \text{für } z > 1 \end{cases}$$

verstanden, so kann statt (3.2) auch

$$(3.12) \qquad M(r,x,f,q) = \frac{2}{\omega B(m,q) r^n} \int_{E_n} f(y) (1-|y-x|^2/r^2)_*^{q-1} dy$$

geschrieben und auf dieses Integral der Hilfssatz angewandt werden. Halten wir $x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n$ und y fest, so ist $(1-|y-x|^2/r^2)_*^{q-1}$ als Funktion von $t = x_i$ totalstetig für $q > 1$ und deshalb

$$(3.13) \qquad \frac{\partial}{\partial x_i} M(r,x,f,q) = \frac{4(q-1)}{\omega B(m,q) r^{n+2}} \int_{E_n} f(y) (y_i - x_i) (1-|y-x|^2/r^2)_*^{q-2} dy,$$

falls auch die übrigen Voraussetzungen des Hilfssatzes erfüllt sind. Dies ist bei stetigem $f(x)$ für $q > 1$, bei lokalsummierbarem $f(x)$ für $q \geq 2$ der Fall [im letzteren Fall dürfen Grenzübergänge $x \to x_0$ in (3.13) unter dem Integralzeichen vorgenommen werden, da $(1-z)_*^{q-2} \leq 1$ ist]. Man erhält so den

Satz 3.2. *Ist $f(x)$ stetig bzw. lokalsummierbar und $q \geq i$ bzw. $q \geq i+1$ $(i = 0, 1, \ldots)$, so ist $M(r,x,f,q)$ als Funktion von x (bei festem $r > 0$) i-mal stetig differenzierbar. Die Formel*

$$(3.14) \qquad D^i M(r,x,f,q) = \frac{2}{\omega B(m,q) r^n} \int_{|y-x|<r} f(y) D^i (1-|y-x|^2/r^2)^{q-1} dy$$

(vgl. zur Bezeichnung Satz 3.1) ist immer dann gültig, wenn f stetig und $q > j$ oder wenn $f \in L_0$ und $q \geq j+1$ ist.

Dieser Satz ist im wesentlichen durch die vorangehenden Bemerkungen bewiesen. Lediglich der Fall, daß f stetig und $q = i$ ist, bedarf noch einer besonderen Überlegung. Es darf dann zunächst gemäß (3.14) $(i-1)$-mal unter dem Integralzeichen differenziert werden. Da aber $D^{i-1}(1-|y-x|^2/r^2)^{i-1}$ ein Polynom in den x_i, y_i ist, entstehen so endlich viele Ausdrücke der Form $P(x) M(r, x, Q(y) f(y), 1)$ (P, Q Polynome), und es bleibt nachzuweisen, daß für stetiges f der Mittelwert $M(r,x,f,1)$ — das ist der gewöhnliche über das Kugelinnere erstreckte Mittelwert — stetig differenzierbar ist. Das ist bekanntlich richtig, und es gilt die Formel (GHERMANESCU [5, S. 290])

$$(3.15) \qquad \begin{cases} \dfrac{\partial}{\partial x_i} M(r,x,f,1) = \dfrac{n}{\omega r^{n+1}} \displaystyle\int_{|y-x|=r} f(y) (y_i - x_i) dy \\ \qquad\qquad = \dfrac{n}{r^2} [M(r,x,y_i f) - x_i M(r,x,f)]. \end{cases}$$

Man kann weitere Sätze dieser Art aufstellen unter der Annahme, daß f lokal quadratisch integrierbar oder lokal von der Klasse L_p, $p > 1$, ist. Doch soll hierauf nicht eingegangen werden. Vielmehr wenden wir uns jetzt den nach r bzw. s genommenen Ableitungen zu.

Satz 3.3. *Ist $f(x)$ stetig und $q \geq 0$ bzw. lokalsummierbar und $q \geq 1$, so gilt für jedes $\alpha > 0$ die Identität*

$$(3.16)\quad \overline{M}(s,x,f,q+\alpha) = \frac{1}{B(m+q,\alpha)\, s^{m+\alpha+q-1}} \int_0^s z^{m+q-1}(s-z)^{\alpha-1}\overline{M}(z,x,f,q)\,dz.$$

Beweis. Es sei zunächst f stetig und $q>0$ [für $q=0$ ist (3.16) mit (3.5) identisch]. Dann ist nach (3.5) (die Argumente x, f sind weggelassen)

$$\begin{aligned}
&\int_0^s z^{m+q-1}(s-z)^{\alpha-1}\overline{M}(z,q)\,dz \\
&\quad = \frac{1}{B(m,q)} \int_0^s (s-z)^{\alpha-1} \int_0^z \sigma^{m-1}\overline{M}(\sigma,0)\,(z-\sigma)^{q-1}\,d\sigma\,dz \\
&\quad = \frac{1}{B(m,q)} \int_0^s \sigma^{m-1}\overline{M}(\sigma,0) \int_\sigma^s (s-z)^{\alpha-1}(z-\sigma)^{q-1}\,dz\,d\sigma \\
&\quad = \frac{B(\alpha,q)}{B(m,q)} \int_0^s \sigma^{m-1}(s-\sigma)^{\alpha+q-1}\overline{M}(\sigma,0)\,d\sigma.
\end{aligned}$$

Hieraus und aus der Identität

$$B(m,q)\,B(m+q,\alpha) = B(\alpha,q)\,B(m,q+\alpha)$$

folgt (3.16).

Nun sei $f \in L_0$ und $q \geq 1$. Dann sind die Funktionen $z^m\overline{M}(z,x,f,q)$ und $z^m\overline{M}(z,x,|f|,q)$ für $z \geq 0$ stetig, die Funktionen $z^{m-1}M(z,x,f,0)$ und $z^{m-1}M(z,x,|f|,0)$ nach z integrierbar. Deshalb sind die gemachten Integralumformungen auch im jetzt vorliegenden Fall erlaubt.

Das in (3.16) rechts stehende Integral ist, wenn $z^{m+q-1}\overline{M}(z,x,f,q)$ für $z \geq 0$ stetig ist, $[\alpha]$-mal[3]) nach s differenzierbar. Aus dieser Bemerkung erhält man leicht den

Satz 3.4. *Ist $f(x)$ stetig bzw. lokalsummierbar, so ist der Mittelwert $\overline{M}(s,x,f,q)$ für $s>0$ $[q]$-mal bzw. $[q-1]$-mal nach s stetig differenzierbar.*

So ergibt sich z. B., wenn in (3.16) $\alpha = 1$ gesetzt und q durch $q-1$ ersetzt wird, durch Differentiation

$$(3.17)\quad s\overline{M}_s(s,x,f,q) = (m+q-1)\,[\overline{M}(s,x,f,q-1) - \overline{M}(s,x,f,q)],$$

gültig für $f \in C^0$, $q \geq 1$ oder für $f \in L_0$, $q \geq 2$. Differenziert man diese Gleichung nochmals, so erhält man nach einfacher Zwischenrechnung

$$(3.18)\quad \begin{cases} s^2\overline{M}_{ss}(s,x,f,q) + (m+q)\,s\overline{M}_s(s,x,f,q) = (m+q-1)\,s\overline{M}_s(s,x,f,q-1) \\ \qquad = (m+q-1)(m+q-2)\,[\overline{M}(s,x,f,q-2) - \overline{M}(s,x,f,q-1)]; \end{cases}$$

[3]) $[\alpha]$ bedeutet die größte ganze Zahl $\leq \alpha$.

dabei ist $f \in C^0$, $q \geq 2$ oder $f \in L_0$, $q \geq 3$ vorausgesetzt. Andererseits folgt aus (3.14), da

$$(3.19)\quad \frac{\Delta(1-|x|^2/s)^{q-1}}{B(m,q)} = \frac{4(m+q-1)(m+q-2)}{s}\left[\frac{(1-|x|^2/s)^{q-3}}{B(m,q-2)} - \frac{(1-|x|^2/s)^{q-2}}{B(m,q-1)}\right]$$

ist (das Ergebnis bleibt richtig, wenn x durch $y-x$ ersetzt wird), die Formel

$$(3.20)\quad \begin{cases} \Delta \overline{M}(s,x,f,q) \\ = \dfrac{4(m+q-1)(m+q-2)}{s}\,[\overline{M}(s,x,f,q-2) - \overline{M}(s,x,f,q-1)] \end{cases}$$

für $f \in C^0$, $q > 2$ oder $f \in L_0$, $q \geq 3$. Bei stetigem $f(x)$ ist das auch noch für $q = 2$ richtig; denn zunächst ist nach (3.13)

$$\frac{\partial}{\partial x_i}\overline{M}(s,x,f,2) = \frac{2(m+1)}{s}\,[\overline{M}(s,x,y_i f,1) - x_i \overline{M}(s,x,f,1)],$$

sodann nach (3.15)

$$\Delta \overline{M}(s,x,f,2) = \frac{2(m+1)}{s}\left[-n\overline{M}(s,x,f,1) + \right.$$
$$\left. + \frac{n}{s}\sum_{i=1}^{n}\{\overline{M}(s,x,y_i^2 f) - 2x_i\overline{M}(s,x,y_i f) + x_i^2 \overline{M}(s,x,f)\}\right],$$

woraus wegen $n = 2m$, $s = \sum_{i=1}^{n}(x_i^2 - 2x_i y_i + y_i^2)$ die behauptete Gl. (3.20) folgt.

Es gilt also nach (3.18) und (3.20) der

Satz 3.5. *Der Mittelwert $\overline{M}(s,x,f,q)$ ist für stetiges f und $q \geq 2$ sowie für lokalsummierbares f und $q \geq 3$ eine im Gebiet $s > 0$ zweimal stetig differenzierbare Lösung der EPD-Gleichung*

$$(3.21)\quad s\overline{M}_{ss}(s,x,f,q) + (m+q)\overline{M}_s(s,x,f,q) - \tfrac{1}{4}\Delta\overline{M}(s,x,f,q) = 0.$$

Die Beziehung

$$(3.22)\qquad \overline{M}(s,x,f,q) \to f(x) \quad \text{für } s \to 0$$

gilt bei stetigem $f(x)$ für alle x [4], bei lokalsummierbarem $f(x)$ für fast alle x, insbesondere in jedem Punkt x, für den $\overline{M}(s,x,f,1) \to f(x)$ strebt für $s \to 0$ [5].

Übrigens folgt aus $\overline{M}(s,x,f,1) \to f(x)$ die Relation (3.22) mit Hilfe von (3.16) sogar für $q \geq 1$.

Satz 3.6. *Ist $f(x)$ stetig differenzierbar, so ist $\overline{M}(s,x,f,q)$ für $q \geq 1$ eine Lösung von (3.21) mit den Randwerten*

$$(3.23)\quad \overline{M}(0,x,f,q) = f(x),\ \overline{M}_{x_i}(0,x,f,q) = f_{x_i}(x),\ \sqrt{s}\,\overline{M}_s(s,x,f,q) \to 0\ (s \to 0)\ [4].$$

[4]) Die Konvergenz ist gleichmäßig, wenn x einen beschränkten Bereich durchläuft.

[5]) Nach einem bekannten Satz von LEBESGUE [7, S. 118] existiert der zuletzt angeschriebene Grenzwert für fast alle x.

Beweis. Nach (3.8) und (3.13) ist für $q>1$

$$\Delta \overline{M}(s,x,f,q)=\frac{4(q-1)}{\omega B(m,q)s^{m+1}}\sum_{i=1}^{n}{}^{E_n}\!\int f_{x_i}(y)(y_i-x_i)(1-|y-x|^2/s)^{q-2}_{*}\,dy,$$

also

$$\text{(3.24)}\quad\begin{cases}\Delta\overline{M}(s,x,f,q)\\ =\dfrac{2(m+q-1)}{s}\displaystyle\sum_{i=1}^{n}[\overline{M}(s,x,y_i f_{x_i},q-1)-x_i\overline{M}(s,x,f_{x_i},q-1)]\end{cases}$$

während aus (3.4)

$$\text{(3.25)}\quad\begin{cases}\overline{M}_s(s,x,f)=\dfrac{1}{2\omega\sqrt{s}}\displaystyle\sum_{i=1}^{n}\int_{|y|=1}f_{x_i}(x+\sqrt{s}\,y)\,y_i\,dy\\ =\dfrac{1}{2s}\displaystyle\sum_{i=1}^{n}[\overline{M}(s,x,y_i f_{x_i})-x_i\overline{M}(s,x,f_{x_i})]\end{cases}$$

und damit aus (3.5)

$$\text{(3.26)}\quad\begin{cases}\overline{M}_s(s,x,f,q-1)=\dfrac{1}{B(m,q-1)}\dfrac{\partial}{\partial s}\displaystyle\int_0^1 t^{m-1}(1-t)^{q-2}\overline{M}(st,x,f)\,dt\\ =\dfrac{1}{B(m,q-1)}\displaystyle\int_0^1 t^{m}(1-t)^{q-2}\frac{1}{2st}\sum_{i=1}^{n}[\overline{M}(st,x,y_i f_{x_i})-x_i\overline{M}(st,x,f_{x_i})]\,dt\\ =\dfrac{1}{2s}\displaystyle\sum_{i=1}^{n}[\overline{M}(s,x,y_i f_{x_i},q-1)-x_i\overline{M}(s,x,f_{x_i},q-1)]\end{cases}$$

folgt. Hieraus, aus (3.24) und aus der ersten Gl. (3.18) (die im vorliegenden Fall offenbar gültig ist) ergibt sich die behauptete Gl. (3.21). Ist $q=1$, so ist (3.24) eine unmittelbare Folgerung aus (3.8) und (3.15), und (3.26) ist mit (3.25) identisch.

Schließlich gilt nach (3.25) für $s\to 0$

$$\sqrt{s}\,M_s(s,x,f)=\frac{1}{2\omega}\sum_{i=1}^{n}\int_{|y|=1}y_i f_{x_i}(x+\sqrt{s}\,y)\,dy\to\frac{1}{2\omega}\sum_{i=1}^{n}f_{x_i}(x)\int_{|y|=1}y_i\,dy=0.$$

Daraus erhält man in einfacher Weise die Limesrelation in (3.23), sogar für alle $q\geq 0$, während die beiden dort stehenden Gleichungen nach (3.6) und (3.8) richtig sind.

Der Vollständigkeit halber führen wir noch einen im wesentlichen bekannten Satz an.

Satz 3.7. *Der Mittelwert $\overline{M}(s,x,f,q)$ ist für zweimal stetig differenzierbares $f(x)$ und $q\geq 0$ eine Lösung von* (3.21) *mit den Randwerten (die Argumente f,q*

sind weggelassen)

$$(3.27)\qquad \begin{cases} \overline{M}(0,x) = f(x), \quad \overline{M}_{x_i}(0,x) = f_{x_i}(x), \quad \overline{M}_{x_i x_j}(0,x) = f_{x_i x_j}(x), \\ \overline{M}_s(0,x) = \dfrac{\Delta f(x)}{4(m+q)}, \quad s\overline{M}_{ss}(s,x) \to 0 \quad (s \to 0)\ ^{6}). \end{cases}$$

Die ersten drei Gln. (3.27) ergeben sich wieder unmittelbar aus (3.6) und (3.8). Aus (3.25) erhält man durch Anwendung des Gaußschen Integralsatzes auf das mittlere Integral

$$(3.28)\qquad \overline{M}_s(s,x,f) = \frac{1}{2n}\overline{M}(s,x,\Delta f,1)$$

und hieraus mit (3.5)

$$\overline{M}_s(s,x,f,q) = \frac{1}{B(m,q)}\int_0^1 t^{m-1}(1-t)^{q-1}\frac{\partial}{\partial s}\overline{M}(st,x,f)\,dt$$

$$= \frac{1}{2n\,B(m,q)}\int_0^1 t^{m}(1-t)^{q-1}\overline{M}(st,x,\Delta f,1)\,dt,$$

also nach (3.16)

$$(3.29)\qquad \overline{M}_s(s,x,f,q) = \frac{1}{4(m+q)}\overline{M}(s,x,\Delta f,q+1).$$

Daraus folgt sowohl die Differentialgleichung (3.21) [mit Hilfe von (3.17), angewandt auf die rechte Seite] als auch die vierte der Beziehungen (3.27), während die letzte dieser Beziehungen nach den vorangehenden und (3.21) gültig ist.

4. Bemerkungen zum Randverhalten der Lösungen. Bei analytischen Randwerten weiß man hierüber nach (2.5), (2.10) und (2.11) genau Bescheid. Hier soll die Frage, ob bei nichtanalytischen, aber hinreichend oft differenzierbaren Anfangswerten wenigstens ein Teil dieser Gleichungen richtig bleibt, untersucht werden. Die Ausnahmewerte $c = 0, -1, -2, \ldots$ werden nicht in Betracht gezogen.

Die angewandte Methode rührt im wesentlichen von WEINSTEIN [9] her. In einer späteren Arbeit [10] gab WEINSTEIN einen anderen Existenzbeweis für das Anfangswertproblem, der weniger Regularitätsvoraussetzungen der Anfangswerte benötigt, bei dem aber das Verhalten der Lösungen in der Nähe von $s = 0$ nicht unmittelbar ersichtlich ist.

Satz 4.1. *Die Funktion $\bar{u}(s,x)$ sei eine Lösung $u^{(c+i)}$ (i natürliche Zahl) von (1.3) (mit $c+i$ an Stelle von c), welche stetige Randwerte $\bar{u}(0,x) = g(x)$ annimmt. Ferner mögen die Ableitungen $\bar{u}_s, \bar{u}_{x_i}, \bar{u}_{x_i x_i}$ für kleine s dem Betrage nach $\leq C s^{\alpha-1}$, $\alpha > 0$, sein*[6]*). Dann ist die Funktion*

$$(4.1)\qquad \bar{w}(s,x) = \int_0^s (s-\sigma)^{i-1}\,\bar{u}(\sigma,x)\,d\sigma$$

[6]) C und α sollen, wenn x einen beschränkten Bereich durchläuft, unabhängig von x gewählt werden können.

eine Lösung der Differentialgleichung

$$s\,\overline{w}_{ss} + c\,w_s - \tfrac{1}{4}\,\Delta\,\overline{w} = (c+i-1)\,s^{i-1}\,g(x)\,. \tag{4.2}$$

Beweis. Da nach den Voraussetzungen unter dem Integralzeichen differenziert werden darf, ist für $i \geq 3$

$$\tfrac{1}{4}\,\Delta\,\overline{w}(s,x) = \tfrac{1}{4}\int_0^s (s-\sigma)^{i-1}\,\Delta\,\bar{u}(\sigma,x)\,d\sigma = \int_0^s (s-\sigma)^{i-1}\,[\sigma\bar{u}_\sigma + (c+i-1)\,\bar{u}]_\sigma\,d\sigma$$

$$= (s-\sigma)^{i-1}\,(\sigma\bar{u}_\sigma + (c+i-1)\,\bar{u})\,|_0^s + (i-1)\int_0^s (s-\sigma)^{i-2}\,[\sigma\bar{u}_\sigma + (c+i-1)\,\bar{u}]\,d\sigma$$

$$= -(c+i-1)\,s^{i-1}\,g(x) + (c+i-1)\,\overline{w}_s(s,x) + (i-1)\,(s-\sigma)^{i-2}\,\sigma\bar{u}\,|_0^s +$$

$$+ (i-1)\int_0^s \bar{u}\,[(i-2)\,s\,(s-\sigma)^{i-3} - (i-1)\,(s-\sigma)^{i-2}]\,d\sigma$$

$$= -(c+i-1)\,s^{i-1}\,g(x) + (c+i-1)\,\overline{w}_s + s\,\overline{w}_{ss} - (i-1)\,\overline{w}_s\,,$$

woraus die Behauptung folgt. Für $i=1$ und $i=2$ sind im Beweis kleine Änderungen anzubringen.

Satz 4.2. *Ist i eine natürliche Zahl und $f(x) \in C^{2i+2}$, so ist*

$$\bar{u}(s,x) = \sum_{\nu=0}^{i-1} \frac{s^\nu\,\Delta^\nu f(x)}{4^\nu\,\nu!\,(c)_\nu} + \frac{1}{4^i\,(i-1)!\,(c)_i}\int_0^s (s-\sigma)^{i-1}\,\overline{M}(\sigma,x,\Delta^i f, c+i-m)\,d\sigma \tag{4.3}$$

für $c+i \geq m$, $c \neq 0, -1, \ldots$, eine Lösung des Anfangswertproblems (1.3), (1.4). *Sie nimmt die Randwerte*

$$\frac{\partial^j\,\bar{u}(0,x)}{\partial s^j} = \frac{\Delta^j f(x)}{4^j\,(c)_j} \qquad \textit{für } j = 0, 1, \ldots, i+1 \tag{4.4}$$

an.

Der Beweis ist sehr einfach. Wird der Operator $\overline{L}$ auf das Integral bzw. auf die endliche Summe in (4.3) angewandt, so ergibt sich beide Male [für das Integral nach (4.2), für die Summe nach dem Beweis zu (2.2)] der Wert $s^{i-1}\,\Delta^i f(x)/[4^i\,(i-1)!\,(c)_{i-1}]$, jedoch mit umgekehrten Vorzeichen. Die Gl. (4.4) ist für $j<i$ unmittelbar klar, für $j=i$ ist nach (4.3)

$$\frac{\partial^i\,\bar{u}(s,x)}{\partial s^i} = \frac{1}{4^i\,(c)_i}\,\overline{M}(s,x,\Delta^i f, c+i-m)\,,$$

für $j=i+1$ wende man (3.29) (mit $\Delta^i f$ an Stelle von f) auf die rechte Seite dieser Gleichung an.

Die Sätze 2.1 und 4.2 stehen in enger Beziehung zueinander. Der letztere ist gewissermaßen eine Verallgemeinerung des ersten auf nichtanalytische Anfangswerte, während man diesen wiederum aus dem bei analytischen Anfangswerten für jedes hinreichend große i gültigen Satz 4.2 hervorgegangen denken kann [in der Tat ist es möglich, die Formel (2.2) aus (4.3) herzuleiten].

5. Bemerkungen zum Eindeutigkeitssatz und zur Aufgabenstellung. Für $k \geq 0$ besteht der folgende Satz ([1, S. 330], [3, S. 381]): Ist $u(r, x) = u^{(k)}$ eine Lösung von (1.1) und gilt $u(0, x) = 0$, $u_r(r, x) \to 0$, $u_{x_i}(r, x) \to 0$ $(r \to 0;$ $i = 1, \ldots, n)$ [7]), so verschwindet u identisch.

Nach diesem Satz ist eine Lösung $u(r, x)$ von (1.1) für $k \geq 0$ durch die vier Forderungen

(α) u ist zweimal stetig differenzierbar für $r > 0$;
(β) u ist stetig differenzierbar für $r \geq 0$;
(γ) $u(0, x) = f(x)$;
(δ) $u_r(0, x) = 0$;

eindeutig bestimmt. (β) und (γ) sind nur für $f \in C^1$ miteinander verträglich, d.h. die Eindeutigkeit der Lösung ist nur für stetig differenzierbare Anfangswerte gesichert. Die Bedingungen (β), (γ), (δ) lassen sich auch durch andere ersetzen (vgl. dazu die erste Fußnote), etwa durch

(β') u ist stetig für $r \geq 0$;
(γ') $u(0, x) = f(x) \in C^1$;
(δ') gleichmäßig in jedem beschränkten Gebiet des E_n gilt $u_{x_i}(r, x) \to f_{x_i}(x)$, $u_r(r, x) \to 0 \quad (r \to 0)$.

Durch eine Abänderung des üblichen Eindeutigkeitsbeweises können diese Aussagen verschärft werden (vgl. zum folgenden [3, S. 381–382]). Es sei ξ ein fester Punkt des E_n und $|x - \xi| = R - r$ $(R > 0)$ ein charakteristischer Kegel. Wir integrieren die für jedes reelle α und jede Lösung $u = u^{(k)}$ gültige Identität

$$2r^\alpha u_r L(u) = \left[r^\alpha \left(u_r^2 + \sum_{i=1}^n u_{x_i}^2\right)\right]_r - 2\sum_i (r^\alpha u_{x_i} u_r)_{x_i} + r^{\alpha-1}\left[(2k - \alpha) u_r^2 - \alpha \sum_i u_{x_i}^2\right] = 0$$

über den Kegelstumpf G: $|x - \xi| \leq R - r$, $0 < r_1 \leq r \leq r_2 < R$; mit den Abkürzungen

$$(5.1) \qquad \begin{cases} E(r) = \int\limits_{|x-\xi|<R-r} u_r^2(r, x)\, dx, \quad F(r) = \int\limits_{|x-\xi|<R-r} \sum\limits_{i=1}^n u_{x_i}^2(r, x)\, dx, \\ G(r) = E(r) + F(r) \end{cases}$$

erhalten wir durch Anwendung des Gaußschen Integralsatzes

$$Q = \int_{r_1}^{r_2} r^{\alpha-1} [2k E(r) - \alpha G(r)]\, dr + Q,$$

---

[7]) Die letzte Bedingung wird a.a.O. nicht ausdrücklich formuliert, aber beim Beweis benötigt. Macht man die zusätzliche Voraussetzung, daß $u \in C^1$ für $r \geq 0$ sein soll, so ist sie natürlich von selbst erfüllt. Es sei aber erwähnt, daß für eine für $r > 0$ zweimal stetig differenzierbare Funktion $v(r, x)$ mit $v(0, x) = 0$, $v_r(r, x) \to 0$ $(r \to 0)$ nicht notwendig $v_{x_i}(r, x) \to 0$ strebt. Gegenbeispiel:

$$v(r, x) = \frac{1}{i^3} \sin(i^4 x_1)\, [\sin(i(i+1)r)]^3 \quad \text{für} \quad \frac{\pi}{i+1} \leq r \leq \frac{\pi}{i}, \quad i = 1, 2, \ldots.$$

wo Q ein über die Oberfläche von G zu erstreckendes Flächenintegral ist, dessen Integrand gleich dem skalaren Produkt aus dem Vektor

$$r^\alpha \left(u_r^2 + \sum_i u_{x_i}^2, -2u_{x_1}u_r, \ldots, -2u_{x_n}u_r\right)$$

und dem nach außen weisenden Normalenvektor $(r^\nu, x_1^\nu, \ldots, x_n^\nu)$ ist. Die der Ober- und Unterseite von G entsprechenden Anteile von Q sind gerade $r_2^\alpha G(r_2)$ und $-r_1^\alpha G(r_1)$, während auf dem Kegelmantel $\sum_i (x_i^\nu)^2 = (r^\nu)^2$ und $r^\nu = \sqrt{2}/2$, also der Integrand gleich

$$\frac{1}{r^\nu} \sum_{i=1}^{n} r^\alpha (u_r x_i^\nu - u_{x_i} r^\nu)^2 \geq 0 \tag{5.2}$$

ist. Bezeichnen wir das Integral über den Kegelmantel mit $P(r_1, r_2) \geq 0$, so ist also

$$\int_{r_1}^{r_2} r^{\alpha-1} [2kE(r) - \alpha G(r)]\, dr + r_2^\alpha G(r_2) - r_1^\alpha G(r_1) + P(r_1, r_2) = 0. \tag{5.3}$$

Wenn dann $\alpha > 0$ ist und u den Bedingungen (α), (β), (γ) genügt, so folgt hieraus für $r_1 \to 0$

$$G(r_2) + r_2^{-\alpha} P(0, r_2) + r_2^{-\alpha} \int_0^{r_2} r^{\alpha-1} [2kE(r) - \alpha G(r)]\, dr = 0. \tag{5.4}$$

Führen wir hierin den Grenzübergang $r_2 \to 0$ durch, so strebt der zweite Summand gegen Null, da das $r_2^{-\alpha}$-fache des Integranden nach (5.2) beschränkt ist und der Inhalt des Integrationsbereiches (nämlich der Mantelfläche) gegen 0 geht. So ergibt sich aus (5.4)

$$G(0) + \frac{2k}{\alpha} E(0) - G(0) = \frac{2k}{\alpha} E(0) = 0.$$

Damit haben wir das bemerkenswerte Ergebnis:

Für $k \neq 0$ ist die Bedingung (δ) von selbst erfüllt. Oder anders gesagt: Bei der Formulierung des Anfangswertproblems kann die Bedingung (δ) weggelassen werden.

Bei der Wellengleichung ($k = 0$) wird neben (1.1), (1.2) noch ein anderes Anfangswertproblem betrachtet, bei dem auf dem Rande ($r = 0$) u verschwinden und die Ableitung u_r vorgeschriebene Werte annehmen soll. Dieses Problem ist für $k \neq 0$ — so kann man das obige Ergebnis auch aussprechen — sinnlos, d.h. es läßt keine Lösung zu [jedenfalls keine den Bedingungen (α) und (β) genügende].

Aus (5.3) kann noch mehr gefolgert werden. Ist $k > 0$ und die Lösung u so beschaffen, daß neben (α), (β'), (γ') und der ersten Bedingung in (δ') noch $r^k u_r(r, x) \to 0$ $(r \to 0)$ [8]) gilt, so erhält man aus (5.3) mit $\alpha = 2k$ für $r_1 \to 0$

$$r_2^\alpha G(r_2) \leq \alpha \int_0^{r_2} r^{\alpha-1} F(r)\, dr = r_2^\alpha F(r^*) \quad \text{mit } 0 \leq r^* \leq r_2,$$

[8]) Der Limes soll gleichmäßig in jedem beschränkten x-Gebiet existieren.

woraus $E(r) \to 0$ $(r \to 0)$ folgt. Strebt darüber hinaus auch $F(r) \to 0$ für $r \to 0$, so kann man in (5.3) jetzt $\alpha = 0$ setzen und $r_1 \to 0$ streben lassen. Da dann auf der linken Seite von (5.3) nur nichtnegative Glieder stehen, ergibt sich der Eindeutigkeitssatz in folgender verschärfter Form.

Eindeutigkeitssatz. *Ist $k > 0$ und $u(r, x)$ eine Lösung der EPD-Gleichung, die für $r \to 0$ zusammen mit den Ableitungen u_{x_i} verschwindet, während für die Ableitung u_r die Relation $r^k u_r(r, x) \to 0$ $(r \to 0)$ gilt*[8]*, dann ist $u = 0$ für $r \geq 0$.*

Die Bedingung $r^k u_r \to 0$ kann, wie das folgende Beispiel zeigt, nicht abgeschwächt werden. Ist nämlich $g(x)$ eine ganze harmonische Funktion, so ist $u = r^{1-k} g(x)$ eine Lösung von (1.1). Diese verschwindet, falls $0 < k < 1$ ist, für $r = 0$ zusammen mit ihren Ableitungen u_{x_i}, während die Ableitung u_r sich für $r \to 0$ wie r^{-k} verhält.

Ein weiteres interessantes Ergebnis erhält man aus (5.3), wenn man diese Formel nach r_2 differenziert:

$$2k r^{\alpha-1} E(r) + r^{\alpha} G'(r) + \frac{\partial}{\partial r} P(r_1, r) = 0 \qquad (r > r_1).$$

Das Integral $P(r_1, r)$ ist mit r monoton wachsend (für jedes feste r_1), da der Integrand nichtnegativ ist und das Integrationsgebiet größer wird. Da außerdem $E(r) \geq 0$ ist, ergibt sich:

Ist $u(r, x)$ irgendeine Lösung der EPD-Gl. (1.1) *und $k \geq 0$, so hat für jeden charakteristischen Kegel die durch* (5.1) *definierte Funktion $G(r)$ im Intervall $0 < r < R$ eine Ableitung $G'(r) \leq 0$.*

Es sei zum Schluß bemerkt, daß hieraus sofort der zu Anfang der Nummer formulierte Eindeutigkeitssatz folgt.

Literatur

[1] ASGEIRSSON, L.: Über eine Mittelwerteigenschaft von Lösungen homogener linearer partieller Differentialgleichungen 2. Ordnung mit konstanten Koeffizienten. Math. Ann. **113**, 321—346 (1937). — [2] BLUM, E. K.: The Euler-Poisson-Darboux-equation in the exceptional case. Proc. Amer. Math. Soc. **5**, 511—520 (1954). — [3] COURANT, R., u. D. HILBERT: Methoden der mathematischen Physik, Bd. 2. Berlin 1937. — [4] DIAZ, J. B., and H. F. WEINBERGER: A solution of the singular initial value problem for the Euler-Poisson-Darboux-equation. Proc. Amer. Math. Soc. **4**, 703—715 (1953). — [5] GHERMANESCU, M.: Sur les valeurs moyennes des fonctions. Math. Ann. **119**, 288—320 (1944). — [6] KAMKE, E.: Das Lebesgue-Stieltjes-Integral. Leipzig 1956. — [7] SAKS, S.: Theory of the integral. Warschau 1937. — [8] WALTER, W.: Mittelwertsätze und ihre Verwendung zur Lösung von Randwertaufgaben. Jahresber. DMV **59**, 93—131 (1957). — [9] WEINSTEIN, A.: Sur le problème de Cauchy pour l'equation de Poisson et l'equation des ondes. C. R. Acad. Sci. Paris **234**, 2584—2585 (1952). — [10] WEINSTEIN, A.: On the wave equation and the equation of Euler-Poisson. Proceedings of Symposia in Applied Mathematics, Vol. V. 1954. — [11] WEINSTEIN, A.: Elliptic and hyperbolic axially symmetric problems. Proceedings of the International Congress of Mathematicians 1954. Vol. III. Amsterdam 1956. S. 264—269.

Mathematisches Institut der Universität Tübingen

(Eingegangen am 12. Dezember 1956)

COLLINGWOOD, E. F.
Math. Zeitschr. Bd. 67, S. 377—396 (1957)

On sets of maximum indetermination of analytic functions

By

E. F. COLLINGWOOD

Dedicated to the memory of GEORGES VALIRON

Introduction

1. The notion of a set of maximum indetermination of a function at a frontier point or set has given rise to significant theorems of a new type in the theory of cluster sets. The purpose of this paper is to develop in further detail the investigation of sets of maximum indetermination and their bearing upon the general theory begun in recent notes[1]). Although the ideas and methods are applicable to other classes of continuous functions, such for example as light interior mappings[2]), and to more general domains, we shall restrict ourselves for the most part to functions meromorphic in the unit circle since this is both the most familiar and the most important case.

We thus take $w = f(z)$ to be meromorphic in $|z| < 1$. In the usual notation[3]) $C(f)$ denotes the cluster set of $f(z)$ in the large, i.e. with respect to the whole circumference $\varkappa \{z : |z| = 1\}$, defined as follows: $a \in C(f)$ if there is a sequence $\{z_n\}$, $|z_n| < 1$, $\lim_{n\to\infty} |z_n| = 1$, such that $\lim_{n\to\infty} f(z_n) = a$. Similarly, $C(f, e^{i\theta})$ is the cluster set of $f(z)$ at the point $z = e^{i\theta}$ defined by substituting the condition $\lim_{n\to\infty} z_n = e^{i\theta}$ for the condition $\lim_{n\to\infty} |z_n| = 1$ above. If $\mathcal{G}$ is a sub-set of the open disc $|z| < 1$ whose closure $\overline{\mathcal{G}}$ intersects $\varkappa$ we define the partial cluster set $C_{\mathcal{G}}(f)$ of $f(z)$ on $\mathcal{G}$ by adding the condition $z_n \in \mathcal{G}$ to the conditions on $\{z_n\}$ in the definition of $C(f)$. Similarly, if $\overline{\mathcal{G}}$ contains $e^{i\theta}$ we define $C_{\mathcal{G}}(f, e^{i\theta})$ by adding the condition $z_n \in \mathcal{G}$ to those in the definition of $C(f, e^{i\theta})$. It is convenient to have a conventional notation for certain special sets $\mathcal{G}$. Thus, if $\mathcal{G}$ is the radius to a point $e^{i\theta}$ we write $C_\varrho(f, e^{i\theta})$ for the (radial) cluster set upon it; and we write $C_{\varrho(\varphi)}(f, e^{i\theta})$ for the cluster set upon the chord to $e^{i\theta}$ making the angle φ, $-\frac{\pi}{2} < \varphi < \frac{\pi}{2}$, with the radius to that point. More generally, we shall use λ_θ to denote a curve terminating at $e^{i\theta}$ and derived from a curve λ_0 terminating at $z = 1$ by rotating it about the origin through

[1]) COLLINGWOOD (2), (3) and (5). Numbers after authors' names refer to the list of references at the end.

[2]) For the definition see WHYBURN (1). STOILOW (1), to whom the systematic development of the theory is due, calls such a mapping simply an interior mapping.

[3]) The general reference for definitions and notation in the theory of cluster sets is to COLLINGWOOD and CARTWRIGHT (1) and COLLINGWOOD (1).

the angle ϑ, $0 \leq \vartheta \leq 2\pi$. σ will denote a spiral path converging to $\varkappa$. The cluster sets upon these curves are denoted by $C_{\lambda_\vartheta}(f, e^{i\vartheta})$ and $C_\sigma(f)$ respectively. A monotonic path λ_ϑ or σ is one which intersects every circle $|z| = r < 1$ in at most one point. A Stolz angle at $e^{i\vartheta}$ will be denoted by Δ and the cluster set in it by $C_\Delta(f, e^{i\vartheta})$. The set $\mathcal{G}$ need not, however, be a continuum although for our purposes it is subject to a density condition. Other notations will be introduced as required.

2. Evidently, $C_\mathcal{G}(f) \subseteqq C(f)$ and $C_\mathcal{G}(f, e^{i\vartheta}) \subseteqq C(f, e^{i\vartheta})$. A set $\mathcal{G}$ is said to be *a set of maximum indetermination of $f(z)$ in the large* if $C_\mathcal{G}(f) = C(f)$, or *of maximum indetermination at $e^{i\vartheta}$* if $C_\mathcal{G}(f, e^{i\vartheta}) = C(f, e^{i\vartheta})$; and the sets $C_\mathcal{G}(f)$ and $C_\mathcal{G}(f, e^{i\vartheta})$ in these cases are *maximal*. The notion is a relative one and can clearly be extended to any two sets $\mathcal{G}_1$, and $\mathcal{G}_2 < \mathcal{G}_1$, where both have limit points on $\varkappa$. The following theorem which generalises an earlier theorem on total cluster sets[4]), is typical. It has been proved[5]) for λ_ϑ monotonic.

Theorem A. *If $f(z)$ is meromorphic in $|z| < 1$, then*

$$C_{\lambda_\vartheta}(f, e^{i\vartheta}) = C(f, e^{i\vartheta})$$

for all ϑ belonging to a residual sub-set of the closed interval $(0, 2\pi)$.

In particular, the radial cluster set $C_\varrho(f, e^{i\vartheta})$ is maximal in a residual set of points of $\varkappa$, a result which supplements and, in a sense, contrasts with FATOU's theorem in the case of a bounded function. For, by that theorem, $C_\varrho(f, e^{i\vartheta})$ is degenerate, that is consists of a single point, almost everywhere on $\varkappa$. But if $C(f, e^{i\vartheta})$ is non-degenerate (and therefore is a continuum) at every point of $\varkappa$ it follows from Theorem A that $C_\varrho(f, e^{i\vartheta})$ is non-degenerate in a residual set and hence that the set of Fatou points $F(f)$ is of the first category. A specific example which illustrates the significance of the theorem is furnished by the Blaschke product

$$B(z) = \prod_{\nu=1}^{\infty} \frac{z e^{-i\vartheta_\nu} - |\zeta_\nu|}{1 - z\bar{\zeta}_\nu},$$

where

$$|\zeta_\nu| < 1, \quad \arg \zeta_\nu = \vartheta_\nu, \quad \sum_{\nu=1}^{\infty} (1 - |\zeta_\nu|) < \infty$$

and the points ζ_ν are so distributed that the closure of their union contains $\varkappa$. Then[6]), for all ϑ, $C(B, e^{i\vartheta}) = \Omega\{w : |w| \leq 1\}$ and so, by Theorem A, $C_\varrho(f, e^{i\vartheta}) = \Omega$ in a residual set of points of $\varkappa$. Another consequence of Theorem A, applied to univalent functions, is that the prime ends of a simply connected domain which contain subsidiary points are a sub-set of category I of the set of prime

[4]) COLLINGWOOD (1) and (4).

[5]) COLLINGWOOD (2).

[6]) This carries further an observation of EGGLESTON (1). See also BAGEMIHL and SEIDEL (1), p. 1070.

ends of the domain. This gives the complete solution of a well known problem of CARATHEODORY [7]).

3. The present paper is in two parts: Part 1 is concerned with pointwise cluster sets, that is to say cluster sets on a set $\mathscr{G}$ at a limit point $e^{i\vartheta}$ of $\mathscr{G}$. The methods are similar to those of my paper (1). Theorem 1, proved in § 4, generalises Theorem A, the monotonic curve λ_ϑ being replaced by a set $\mathscr{G}_\vartheta$ derived by rotation from an initial set $\mathscr{G}_0$. We next, in §§ 5—6, consider the set $J(f)$ of points $e^{i\vartheta}$ at which, for every Stolz angle Δ, the cluster set $C_\Delta(f, e^{i\vartheta})$ is maximal and the larger set $K(f)$ of points $e^{i\vartheta}$ at which, given any pair of Stolz angles Δ_1 and Δ_2, $C_{\Delta_1}(f, e^{i\vartheta}) = C_{\Delta_2}(f, e^{i\vartheta})$. $J(f)$ and therefore $K(f)$ also are shown to be residual. $K(f)$ is of interest since it contains both the Fatou points and the Plessner points [8]) of $f(z)$. On the same principle we prove in § 7 that the boundary cluster sets at $e^{i\vartheta}$ are also maximal in a residual set on $\varkappa$. In § 8 we extend to points of $K(f)$ a previous theorem on Plessner points and in § 9 we consider the so-called principal values [9]) of $C(f, e^{i\vartheta})$ and show that for certain classes of functions these cover $C(f, e^{i\vartheta})$ at a residual set on $\varkappa$, thus generalising the result on prime ends referred to above.

In Part 2 we consider the cluster set $C_\sigma(f)$ of $f(z)$ on an arbitrary spiral path σ in $|z|<1$ which converges to the circumference $\varkappa$ and we prove a theorem (Theorem 10) from which it follows that for certain classes of functions this is maximal for all such paths, with a weaker result in more general cases. There are two proofs of independent interest, one based on a recent theorem of COLLINGWOOD and CARTWRIGHT and the other on an earlier theorem of GROSS and IVERSEN. These are given in §§ 10—11. The methods of Part 1 are not relevant and there is no theorem analogous to Theorem 1 for a rotational family of spirals valid for all meromorphic functions. We conclude in §§ 12—14 with some remarks on the application of our methods and results to more general domains and classes of functions and we prove the analogues of our main theorems for circle domains of finite connectivity.

Part 1

4. Maximum indetermination on rotations of a given set

Let $\mathscr{G}_0$ be a set in $|z|<1$ such that (i) the set $E\{|z| : z \in \mathscr{G}_0\}$ is dense in an interval $r_0<|z|<1$ and (ii) given $\eta>0$ there is a $\delta>0$ such that $|\arg z|<\eta$ for all $z \in \mathscr{G}_0$ contained in the annulus $1-\delta<|z|<1$. We denote by $\mathscr{G}_\vartheta$ the transform of $\mathscr{G}_0$ by rotation about the origin through the angle ϑ so that $\mathscr{G}_\vartheta$ has $z=e^{i\vartheta}$ as its only limit point on $\varkappa$. We now prove

Theorem 1. *If $f(z)$ is meromorphic in $|z|<1$ and if $\{\mathscr{G}_\vartheta\}$ is the family of rotations of a given set which satisfies conditions* (i) *and* (ii), *then for a residual*

[7]) The proof is given in COLLINGWOOD (6). We write category I and II for first and second category.

[8]) For the definitions see § 6 below and footnote 11.

[9]) The notion of principal value defined in the first place by CARATHEODORY for Prime Ends was extended by GROSS (1) to the cluster sets of multi-valent functions.

set of values of ϑ in $0 \leqq \vartheta \leqq 2\pi$, $\mathscr{G}_\vartheta$ is a set of maximum indetermination of $f(z)$ at $e^{i\vartheta}$ i.e.

$$C_{\mathscr{G}_\vartheta}(f, e^{i\vartheta}) = C(f, e^{i\vartheta}). \tag{1}$$

We shall show that the hypothesis $C_{\mathscr{G}_\vartheta}(f, e^{i\vartheta}) \neq C(f, e^{i\vartheta})$ in a set of values of ϑ of category II in $(0, 2\pi)$ leads to a contradiction. The set $C_{\mathscr{G}_\vartheta}(f, e^{i\vartheta})$, like $C(f, e^{i\vartheta})$, is closed and we denote by $C_{\mathscr{G}_\vartheta}(f, e^{i\vartheta})_{+\varepsilon}$ the closed set of points on the unit w-sphere, which we take as the image space of $|z|<1$ by $f(z)$, whose (spherical) distance from $C_{\mathscr{G}_\vartheta}(f, e^{i\vartheta})$ does not exceed ε, where $\varepsilon > 0$. Now since $C_{\mathscr{G}_\vartheta}(f, e^{i\vartheta}) \subseteqq C(f, e^{i\vartheta})$ it follows that if $C_{\mathscr{G}_\vartheta}(f, e^{i\vartheta}) \neq C(f, e^{i\vartheta})$ there is a positive ε such that $C(f, e^{i\vartheta}) \cap \mathscr{C}\left(C_{\mathscr{G}_\vartheta}(f, e^{i\vartheta})_{+\varepsilon}\right)$ is not empty[10]). We choose a decreasing positive sequence $\{\varepsilon_n\}$, $\varepsilon_n \to 0$ as $n \to \infty$. For a given n let E_n be the set of points $z = e^{i\vartheta}$ such that $C(f, e^{i\vartheta}) \cap \mathscr{C}\left(C_{\mathscr{G}_\vartheta}(f, e^{i\vartheta})_{+\varepsilon_n}\right)$ is not empty. Plainly E_n may be empty; and $E_1 \subseteqq E_2 \subseteqq \cdots \subseteqq E_n \subseteqq \cdots$. Denoting by E the set of points $e^{i\vartheta}$ for which $C_{\mathscr{G}_\vartheta}(f, e^{i\vartheta}) \neq C(f, e^{i\vartheta})$ we see that

$$E = \bigcup_{1 \leqq n < \infty} E_n. \tag{2}$$

Now suppose that E is of category II on $\varkappa$. There must then be a number N such that E_N is of category II on $\varkappa$. We now lay down on the unit w-sphere a finite triangular lattice the meshes of which, denoted by $T_1, T_2, \ldots, T_m$, say, are all of (spherical) diameter less than $\varepsilon_N/4$. For a given $\mu \leqq m$ we denote by $E_{N,\mu}$ the sub-set of E_N in which $T_\mu \cap C(f, e^{i\vartheta}) \cap \mathscr{C}\left(C_{\mathscr{G}_\vartheta}(f, e^{i\vartheta})_{+\varepsilon_N}\right)$ is not empty. Any of the sets $E_{N,\mu}$, $\mu = 1, 2, \ldots, m$ may be empty and any two of them may overlap; but since

$$E_N = \bigcup_{\mu \leqq m} E_{N,\mu}$$

we can find $\mu = M$, say, such that $E_{N,M}$ is of category II on $\varkappa$.

We now sub-divide this set $E_{N,M}$ as follows. First choose a point $a \in T_M$. Since T_M is of diameter less than $\varepsilon_N/4$ and also, for $e^{i\vartheta} \in E_{N,M}$, contains some point of $\mathscr{C}\left(C_{\mathscr{G}_\vartheta}(f, e^{i\vartheta})_{+\varepsilon_N}\right)$ it follows that for all $e^{i\vartheta} \in E_{N,M}$ the distance of a from $C_{\mathscr{G}_\vartheta}(f, e^{i\vartheta})$ is greater than $3\varepsilon_N/4$. Now let $\mathscr{E}_1$, be the sub-set of $E_{N,M}$ such that the spherical distance $[f(z), a]$ satisfies

$$[f(z), a] > \varepsilon_N/2 \tag{3}$$

for all $z \in \mathscr{G}_\vartheta$, $e^{i\vartheta} \in \mathscr{E}_1$, $1/2 < |z| < 1$; $\mathscr{E}_2$ the sub-set of $E_{N,M}$ such that (3) is satisfied for all $z \in \mathscr{G}_\vartheta$, $e^{i\vartheta} \in \mathscr{E}_2$, $3/4 < |z| < 1$, and, generally, $\mathscr{E}_p$, $p < \infty$, the sub-set of $E_{N,M}$ such that (3) is satisfied for all $z \in \mathscr{G}_\vartheta$, $e^{i\vartheta} \in \mathscr{E}_p$, $1 - 2^{-p} < |z| < 1$. Evidently any finite number of the sets $\mathscr{E}_p$ may be empty, $\mathscr{E}_1 \subseteqq \mathscr{E}_2 \subseteqq \cdots \subseteqq \mathscr{E}_p \subseteqq \cdots$, and

$$E_{N,M} = \bigcup_{1 \leqq p < \infty} \mathscr{E}_p. \tag{4}$$

[10]) $\mathscr{C}$ denotes the complement of a set with respect to the containing space, in this case the surface of the sphere.

Since $E_{N,M}$ is of category II there is a number P such that $\mathscr{E}_P$ is of category II on $\varkappa$; and there is also an arc β of $\varkappa$ on which $\mathscr{E}_P$ is dense. As $e^{i\vartheta}$ traverses β the set $\mathscr{G}_\vartheta$ sweeps out a set $\mathscr{B}$ of circular arcs. Let β' be an open arc contained in β and let 2η be its angular distance from the two end-points of β. We can find $\delta>0$ such that in the annulus $1-\delta<|z|<1$ all points $z\in\mathscr{G}_\vartheta$ lie in the angle $|\arg z-\vartheta|<\eta$. Consequently the end-points of the arcs of the set $\mathscr{B}$ contained in this annulus $1-\delta<|z|<1$ all lie outside the sector bounded by β' and the radii to its end points; and these arcs of $\mathscr{B}$ therefore traverse the annular sector of width δ bounded by β'. Now choose $\delta'<\min[\delta, 2^{-P}]$ and denote by Σ the annular sector bounded by β' and contained in $1-\delta'<|z|<1$. The set of points $\mathscr{B}\cap\Sigma$ is dense in Σ and for all $z\in\mathscr{B}\cap\Sigma$ the inequality (3) is satisfied. Hence, by continuity,

$$[f(z), a]\geq\varepsilon_N/2$$

for all $z\in\Sigma$. From this it follows, since $a\in T_M$ and T_M is of diameter $<\varepsilon_N/4$, that $T_M\in\mathscr{C}\,C(f, e^{i\vartheta'})$ for all $e^{i\vartheta'}\in\beta'$. But since $\mathscr{E}_P$ is dense on β and therefore also on β' we can find $e^{i\vartheta'}\in\beta'\cap\mathscr{E}_P$ so that, for this value of ϑ',

$$T_M\cap C(f, e^{i\vartheta'})\cap\mathscr{C}\left(C_{\mathscr{G}_{\vartheta'}}(f, e^{i\vartheta'})_{+\varepsilon_N}\right)$$

is not empty and therefore $T_M\cap C(f, e^{i\vartheta'})$ is not empty. This contradiction proves the theorem.

By a boundary path we mean a continuous curve which converges either to a point or an arc of $\varkappa$ which is its end. The end may be the whole circumference. By boundary paths of fixed form to a set of ends we mean a set of curves derived by rotation from a given continuous curve. Given any boundary path λ_0 converging to $z=1$ and the family $\{\lambda_\vartheta\}$ of rotational transforms of λ_0 each converging to $e^{i\vartheta}$ we have, as an immediate consequence of Theorem 1, the following result which contains Theorem A.

Theorem 2. *If $f(z)$ is meromorphic in $|z|<1$ and if $\{\lambda_\vartheta\}$ is any family of boundary paths of fixed form to the corresponding end points $e^{i\vartheta}$, then λ_ϑ is a path of maximum indetermination of $f(z)$ at $e^{i\vartheta}$ in a residual set of points $e^{i\vartheta}$ of the circumference $\varkappa$.*

For λ_ϑ satisfies the conditions on $\mathscr{G}_\vartheta$.

We may remark that if $\mathscr{G}_\vartheta$ lies on a monotonic boundary path λ_ϑ then it follows from continuity that $C_{\mathscr{G}_\vartheta}(f, e^{i\vartheta})=C_{\lambda\vartheta}(f, e^{i\vartheta})$, but this equation does not follow directly from continuity if $\mathscr{G}_\vartheta\subseteqq\lambda_\vartheta$ which is not monotonic. Then, however, $C_{\mathscr{G}_\vartheta}(f, e^{i\vartheta})\subseteqq C_{\lambda_\vartheta}(f, e^{i\vartheta})$ if $\mathscr{G}_\vartheta\subseteqq\lambda_\vartheta$ and it follows from Theorem 1 that $C_{\mathscr{G}_\vartheta}(f, e^{i\vartheta})=C_{\lambda_\vartheta}(f, e^{i\vartheta})=C(f, e^{i\vartheta})$ in a residual set of points $e^{i\vartheta}\in\varkappa$.

Corollary 1. *If $f(z)$ is meromorphic in $|z|<1$ and the sets $\mathscr{G}_\vartheta(1), \mathscr{G}_\vartheta(2), \ldots, \mathscr{G}_\vartheta(n), \ldots$ all satisfy the conditions of Theorem 1 on $\mathscr{G}_\vartheta$, then the set of points $e^{i\vartheta}$ at which every one of the sets $\mathscr{G}_\vartheta(1), \mathscr{G}_\vartheta(2), \ldots, \mathscr{G}_\vartheta(n), \ldots$ is a set of maximum indetermination of $f(z)$ is residual on $\varkappa$.*

For the set of points $e^{i\vartheta}$ in which any one of the sets $\mathscr{G}_\vartheta(n)$, $n=1, 2, \ldots$, is not of maximum indetermination is of category I on $\varkappa$.

5. Let $-\pi/2<\varphi_n<\pi/2$ and let $\mathscr{G}_\vartheta(n)$ be the chord to $e^{i\vartheta}$ making the angle φ_n with the radius vector at this point so that $C_{\mathscr{G}_\vartheta(n)}(f, e^{i\vartheta}) = C_{\varrho(\varphi_n)}(f, e^{i\vartheta})$. With the sets $\mathscr{G}_\vartheta(n)$ specified in this way Corollary 1 gives at once

Corollary 2. *If $f(z)$ is meromorphic in $|z|<1$ and we are given any enumerable set $\varphi_1, \varphi_2, \ldots, \varphi_n, \ldots, -\pi/2<\varphi_n<\pi/2$, then the set of points $e^{i\vartheta}$ for which*

$$C(f, e^{i\vartheta}) = C_{\varrho(\varphi_1)}(f, e^{i\vartheta}) = C_{\varrho(\varphi_2)}(f, e^{i\vartheta}) = \cdots = C_{\varrho(\varphi_n)}(f, e^{i\vartheta}) = \cdots$$

is residual on $\varkappa$.

We shall denote by $J(f)$ *the set of points $e^{i\vartheta}\in\varkappa$ such that for every Stolz angle Δ at $e^{i\vartheta}$ the equation $C_\Delta(f, e^{i\vartheta}) = C(f, e^{i\vartheta})$ is satisfied by $f(z)$*, i.e. every Δ is a set of maximum indetermination of $f(z)$ at every point $e^{i\vartheta}\in J(f)$.

Theorem 3. *If $f(z)$ is meromorphic in $|z|<1$, the set $J(f)$ is a residual set of the circumference $\varkappa$.*

For, $C_\Delta(f, e^{i\vartheta})\subseteqq C(f, e^{i\vartheta})$; and so if we take the set $\{\varphi_n\}$ of Corollary 2 to be a dense enumerable set, say the rationals, in the open interval $(-\pi/2, \pi/2)$, it follows from the Corollary that there is a residual set $\{e^{i\vartheta}\}$, $e^{i\vartheta}\in J(f)$, in which every Stolz angle Δ contains a chord $\varrho(\varphi)$, φ rational, such that

$$C_{\varrho(\varphi)}(f, e^{i\vartheta}) = C(f, e^{i\vartheta});$$

and since

$$C_{\varrho(\varphi)}(f, e^{i\vartheta})\subseteqq C_\Delta(f, e^{i\vartheta})$$

this proves the theorem.

We now introduce *the set $K(f)$ of points $e^{i\vartheta}\in\varkappa$ such that given any two Stolz angles Δ_1, Δ_2 at $e^{i\vartheta}$, whether overlapping or not, the equation $C_{\Delta_1}(f, e^{i\vartheta}) = C_{\Delta_2}(f, e^{i\vartheta})$ is satisfied.* Evidently $J(f)\subseteqq K(f)$ and hence, as an immediate consequence of Theorem 3 we have

Theorem 4. *If $f(z)$ is meromorphic in $|z|<1$ the set $K(f)$ is a residual set of the circumference $\varkappa$.*

6. Certain properties of the sets $J(f)$ and $K(f)$

In the notation and terminology of COLLINGWOOD (1) we define the *outer angular cluster set* $C_{\mathscr{A}}(f, e^{i\vartheta})$ of $f(z)$ at $e^{i\vartheta}$ as the union

$$C_{\mathscr{A}}(f, e^{i\vartheta}) = \bigcup_{\Delta} C_\Delta(f, e^{i\vartheta})$$

taken over all Stolz angles Δ at $e^{i\vartheta}$. In general $C_{\mathscr{A}}(f, e^{i\vartheta})$ is a non-empty F_σ since every $C_\Delta(f, e^{i\vartheta})$ is a non-empty closed set; but for $e^{i\vartheta}\in K(f)$ the set $C_{\mathscr{A}}(f, e^{i\vartheta})$ is itself closed. The set $F(f)$ *of Fatou points of $f(z)$* is the set of points $e^{i\vartheta}$ for which $C_{\mathscr{A}}(f, e^{i\vartheta})$ is degenerate, i.e. consists of a single point. This definition, although it is superficially more general is in fact equivalent to that previously adopted[11]). For if $C_{\mathscr{A}}(f, e^{i\vartheta})$ consists of the point $a(\vartheta)$

[11]) Note that the statement in the final sentence of footnote 3 of COLLINGWOOD (1) is incorrect if $f(z)$ is restricted to be meromorphic.

which we may call *the Fatou value of* $f(z)$ *at* $e^{i\vartheta}$, then given any Stolz angle Δ at $e^{i\vartheta}$ and any number $b \neq a(\vartheta)$ we have $b \in \mathscr{C}\, C_\Delta(f, e^{i\vartheta})$. It follows that given any $\varepsilon < |b - a(\vartheta)|$ we can find δ such that for all z in the angular domain defined by $z \in \Delta$ and $|z - e^{i\vartheta}| < \delta$ the inequality $|f(z) - b| > \varepsilon$ holds; or if, as we may assume without loss of generality, $b = \infty$, then $|f(z)| < 1/\varepsilon$ in the angular domain. It now follows from LINDELÖF's theorem, since $f(z) \to a(\vartheta)$ as $z \to e^{i\vartheta}$ along a chord in Δ, that $f(z)$ tends uniformly to $a(\vartheta)$ in Δ. Since Δ is an arbitrary Stolz angle at $e^{i\vartheta}$ this establishes the Fatou property of uniform convergence in any Stolz angle Δ at $e^{i\vartheta}$.

Now for $e^{i\vartheta} \in K(f)$ the equation

$$C_\Delta(f, e^{i\vartheta}) = C_{\mathscr{A}}(f, e^{i\vartheta})$$

is satisfied for every Stolz angle Δ; and since this equation is also satisfied for $e^{i\vartheta} \in F(f)$ we see that $F(f) \subseteqq K(f)$.

The set $I(f)$ *of Plessner points of* $f(z)$ is the set of points $e^{i\vartheta}$ such that $C_\Delta(f, e^{i\vartheta})$ is total[11a]) for every Stolz angle Δ at $e^{i\vartheta}$. It follows that also $I(f) \subseteqq K(f)$, so that

$$F(f) \cup I(f) \subseteqq K(f).$$

We thus see that the antithetic character of the sets $F(f)$ and $I(f)$ is in a sense only apparent and arises from the degeneracy of $C_\Delta(f, e^{i\vartheta})$ in the set $F(f)$.

By PLESSNER's Theorem[12]), $F(f) \cup I(f)$ is p.p. and therefore we have

Theorem 4a. *For any function* $f(z)$ *meromorphic in* $|z| < 1$ *the set* $K(f)$ *is both residual and* p.p. *on the circumference* $\varkappa$.

We should note that it is not true in general that $F(f) \cup I(f)$ is residual. A bounded function for which $C(f, e^{i\vartheta})$ is not degenerate for any $e^{i\vartheta}$, as is the case when $f(z)$ is a Blaschke product, is a counter example[13]). For then $I(f)$ is empty, and $F(f) \subseteqq \mathscr{C}\, J(f)$, which is of category I by Theorem 3. Thus the set $K(f) \cap \mathscr{C}\, F(f) \cap \mathscr{C}\, I(f)$ may be residual but is always of measure zero. As in previous papers we denote by $W(f)$ the set of points $e^{i\vartheta}$ for which $C(f, e^{i\vartheta})$ is total[14]). Plainly, $W(f) \cap \mathscr{C}\, I(f) \subseteqq \mathscr{C}\, J(f)$, which is of category I. Since $I(f) \subseteqq W(f)$ it follows that $I(f)$ *differs from* $W(f)$ *at most by a set of category I*, which is a known result[15]). Other known results are immediate consequences of Theorems 1, 2 and 3. Here I depart from the notation previously used and write $W_\varrho(f)$ for the set of points $e^{i\vartheta}$ for which the radial cluster set $C_\varrho(f, e^{i\vartheta})$ is total, $W_{\varrho(\varphi)}(f)$ for the set $\{e^{i\vartheta}\}$ for which the chordal cluster set $C_{\varrho(\varphi)}(f, e^{i\vartheta})$ is total and $W_{\lambda\vartheta}(f)$ the set $\{e^{i\vartheta}\}$ for which $C_{\lambda\vartheta}(f, e^{i\vartheta})$ is total, $\{\lambda_\vartheta\}$ being a family of curves of fixed form to the end points $e^{i\vartheta}$. More generally, $W_{\mathscr{I}_\vartheta}(f)$ denotes the set $\{e^{i\vartheta}\}$ for which $C_{\mathscr{I}_\vartheta}(f, e^{i\vartheta})$ is total.

[11a]) A set is total if its complement with respect to the sphere is empty.

[12]) PLESSNER (1) or COLLINGWOOD (1).

[13]) See footnote 6.

[14]) COLLINGWOOD and CARTWRIGHT (1) and COLLINGWOOD (1).

[15]) COLLINGWOOD (4).

Thus the symbol W denotes the (WEIERSTRASS) property of totality of the cluster set indicated. Similarly, we denote by $D(f)$ the set $\{e^{i\theta}\}$ of points, sometimes called convergence points, such that $C(f, e^{i\theta})$ is degenerate; by $D_\varrho(f)$ the set such that $C_\varrho(f, e^{i\theta})$ is degenerate; by $D_{\varrho(\varphi)}(f)$, $D_{\lambda_\theta}(f)$ and $D_{\mathcal{G}_\theta}(f)$ the sets for which $C_{\varrho(\varphi)}(f, e^{i\theta})$, $C_{\lambda_\theta}(f, e^{i\theta})$ and $C_{\mathcal{G}_\theta}(f, e^{i\theta})$ respectively are degenerate. For a bounded function, or a linear transform of a bounded function, we note that $D_\varrho(f) = D_{\varrho(\varphi)}(f) = D_{\lambda_\theta}(f) = F(f)$, but not in the general case. The symbol D denotes the property of degeneracy of the cluster set, or convergence of the function, on the set indicated[16]). With these notations we have

Theorem 5. *If $f(z)$ is meromorphic in $|z|<1$, then any two of the sets $W(f)$, $W_{\mathcal{G}_\theta}(f)$ and $I(f)$ where $\mathcal{G}_\theta$ satisfies the conditions of Theorem 1, differ at most by a set of category I on the circumference $\varkappa$.*

For, by Theorem 1, $W(f) \cap \mathcal{C}\, W_{\mathcal{G}_\theta}(f)$ is of category I and since $W(f) \geqq W_{\mathcal{G}_\theta}(f)$ it follows that $W(f)$ differs from $W_{\mathcal{G}_\theta}(f)$ by a set of category I. We have seen that $W(f)$ differs from $I(f)$ by a set of category I, so that the theorem is proved. If we particularise $\mathcal{G}_\theta$ by making it the radius ϱ or a chord $\varrho(\varphi)$ to $e^{i\theta}$ we obtain a previous theorem of the writer[17]).

Similarly, we have a theorem for the distribution of degenerate cluster sets which, on particularising $\mathcal{G}_\theta$ in the appropriate way, contains previous results of EGGLESTON and the writer[18]).

Theorem 6. *If $f(z)$ is meromorphic in $|z|<1$, then any two of the sets $D(f)$, $D_{\mathcal{G}_\theta}(f)$ and $F(f)$, where the set $\mathcal{G}$ satisfies the conditions of Theorem 1, differ at most by a set of category I on the circumference $\varkappa$.*

For, by Theorem 1, $D_{\mathcal{G}_\theta}(f) \cap \mathcal{C}\, D(f)$ is of category I and since $D(f) \subseteq D_{\mathcal{G}_\theta}(f)$ it follows that $D(f)$ differs from $D_{\mathcal{G}_\theta}(f)$ only by a set of category I. In particular, $D(f)$ differs from $D_\varrho(f)$ only by a set of category I and since, trivially, $D(f) \subseteq F(f) \subseteq D_\varrho(f)$, the theorem is proved.

7. The boundary cluster sets

The set $\mathcal{G}_\theta$ of Theorem 1 is an internal set of $|z|<1$ converging to $e^{i\theta}$ but may be arbitrarily close to the circumference $\varkappa$. In particular, we could take for the sequence $\mathcal{G}_\theta(1), \mathcal{G}_\theta(2), \ldots, \mathcal{G}_\theta(n), \ldots$ of Corollary 1 a sequence of curves $\lambda_\theta(1), \lambda_\theta(2), \ldots, \lambda_\theta(n), \ldots$ to $e^{i\theta}$ in $|z|<1$ and of increasing order of contact with the circumference at $e^{i\theta}$. However, by a slight elaboration of the proof of Theorem 1 we can prove the result which this suggests, namely that *in a residual set of points $e^{i\theta}$ the boundary cluster sets are maximal.* We shall in fact prove a more general theorem analogous to Theorem 1.

Denote by $(\vartheta-\eta, \vartheta)$ and $(\vartheta, \vartheta+\eta)$ the open intervals $\vartheta-\eta<\vartheta'<\vartheta$ and $\vartheta<\vartheta'<\vartheta+\eta$ respectively and let the set $H(\vartheta)$ be dense in one or both of

[16]) In previous papers, EGGLESTON (1) and COLLINGWOOD (5), the set we here call $D(f)$ was denoted by $G(f)$ and the set we here call $D_\varrho(f)$ by $D(f)$.

[17]) COLLINGWOOD (4).

[18]) EGGLESTON (1) and COLLINGWOOD (5).

them. For the moment we assume $H(\vartheta)$ to be dense in both and we denote by $H(\vartheta, -\eta)$ and $H(\vartheta, +\eta)$ the intersections $H(\vartheta) \cap (\vartheta - \eta, \vartheta)$ and $H(\vartheta) \cap (\vartheta, \vartheta + \eta)$. We now write, generalising the familiar definitions of the right and left-hand boundary cluster sets and the complete boundary cluster set of $f(z)$ at $e^{i\vartheta}$,

$$C\left(f, H(\vartheta, -\eta)\right) = \bigcup_{\vartheta' \in H(\vartheta, -\eta)} C(f, e^{i\vartheta'})$$

and

$$(5) \qquad C_{Hr}(f, e^{i\vartheta}) = \bigcap_{\eta} \overline{C}\left(f, H(\vartheta, -\eta)\right); \qquad C\left(f, H(\vartheta, +\eta)\right) = \bigcup_{\vartheta' \in H(\vartheta, +\eta)} C(f, e^{i\vartheta'})$$

and

$$(6) \qquad C_{Hl}(f, e^{i\vartheta}) = \bigcap_{\eta} \overline{C}\left(f, H(\vartheta, +\eta)\right);$$

and

$$(7) \qquad C_H(f, e^{i\vartheta}) = C_{Hr}(f, e^{i\vartheta}) \cup C_{Hl}(f, e^{i\vartheta}).$$

If $H(\vartheta)$ contains an interval $(\vartheta - \eta, \vartheta + \eta)$ then $C_{Hr}(f, e^{i\vartheta})$ and $C_{Hl}(f, e^{i\vartheta})$ become the right and left hand boundary cluster sets $C_{Br}(f, e^{i\vartheta})$ and $C_{Bl}(f, e^{i\vartheta})$ defined by[19]) the equations

$$C_{Br}(f, e^{i\vartheta}) = \bigcap_{\eta} \overline{C}\left(f, (\vartheta - \eta, \vartheta)\right)$$

and

$$C_{Bl}(f, e^{i\vartheta}) = \bigcap_{\eta} \overline{C}\left(f, (\vartheta, \vartheta + \eta)\right),$$

while

$$C_B(f, e^{i\vartheta}) = C_{Br}(f, e^{i\vartheta}) \cup C_{Bl}(f, e^{i\vartheta}).$$

We now define $H(\vartheta)$ as the translation through the interval $\vartheta - \vartheta_0$ of a set $H(\vartheta_0)$, say. Identifying 0 with 2π this gives us a family of sets on the circumference $\varkappa$ which are rotational transforms of one another like the family $\{\mathcal{G}_\phi\}$ of Theorem 1. We proceed to prove

Theorem 7. *If $f(z)$ is meromorphic in $|z| < 1$ and if the family of sets $\{H(\vartheta)\}$ are translations of one another, $H(0)$ being dense in an interval $(-\eta, 0)$, then the equation*

$$(8) \qquad C_{Hr}(f, e^{i\vartheta}) = C(f, e^{i\vartheta})$$

is satisfied for a set of values of ϑ residual in $0 \leq \vartheta \leq 2\pi$. Similarly, if $H(\vartheta)$ is dense in $(\vartheta, \vartheta + \eta)$,

$$(9) \qquad C_{Hl}(f, e^{i\vartheta}) = C(f, e^{i\vartheta})$$

in a residual set in $0 \leq \vartheta \leq 2\pi$.

It is sufficient to prove (8). First, $C_{Hr}(f, e^{i\vartheta}) \subseteq C_{Br}(f, e^{i\vartheta}) \subseteq C(f, e^{i\vartheta})$. Suppose, then, that the set E of values of ϑ for which

$$C_{Hr}(f, e^{i\vartheta}) \neq C(f, e^{i\vartheta})$$

[19]) Notation of COLLINGWOOD and CARTWRIGHT (1), pp. 121—122.

is of category II. Our argument closely follows the proof of Theorem 1 so that a good deal of the detail may be omitted. Given $\varepsilon>0$, $C_{H,r}(f, e^{i\vartheta})_{+\varepsilon}$ is the closed ε-neighbourhood of $C_{H,r}(f, e^{i\vartheta})$ on the sphere. As before, let $\{\varepsilon_n\}$ be a decreasing sequence of positive numbers $\varepsilon_n \to 0$ as $n \to \infty$ and denote by E_n the subset of E for which

$$C(f, e^{i\vartheta}) \cap \mathscr{C}\left(C_{H,r}(f, e^{i\vartheta})_{+\varepsilon_n}\right)$$

is not empty. Then $E_1 \subseteq E_2 \subseteq \cdots \subseteq E_n \subseteq \cdots$, and

$$E = \bigcup_{1 \leq n < \infty} E_n.$$

Hence there exists a number N such that E_N is of category II. We repeat the triangulation $T_1, T_2, \ldots, T_m, T_\mu$ being of (spherical) diameter less than $\varepsilon_N/4$ for all $\mu \leq m$, and we denote by $E_{N,\mu}$ the sub-set of E_N in which

$$T_\mu \cap C(f, e^{i\vartheta}) \cap \mathscr{C}\left(C_{H,r}(f, e^{i\vartheta})_{+\varepsilon_N}\right)$$

is not empty. Since

$$E_N = \bigcup_{\mu \leq m} E_{N,\mu}$$

we can find M such that $E_{N,M}$ is of category II. Now choose $a \in T_M$. The distance of a from $C_{H,r}(f, e^{i\vartheta})$ evidently exceeds $3\varepsilon_N/4$. It follows from (5) that, for $\vartheta \in E_{N,M}$, there is an $\eta = \eta(\vartheta)$ such that

$$a \in \mathscr{C}\left(\overline{C}(f, H(\vartheta, -\eta))_{+\varepsilon_N/2}\right).$$

Choose a sequence $\{\eta_l\}$, $\eta_1 > \eta_2 > \cdots > \eta_l > \cdots$, $\lim_{l \to \infty} \eta_l = 0$, and denote by $E_{N,M,l}$ the sub-set of $E_{N,M}$ for which

$$a \in \mathscr{C}\left(\overline{C}(f, H(\vartheta, -\eta_l))_{+\varepsilon_N/2}\right).$$

Plainly $E_{N,M,l_1} \subseteq E_{N,M,l_2}$ where $l_1 < l_2$, and

$$E_{N,M} = \bigcup_l E_{N,M,l}.$$

Since $E_{N,M}$ is of category II we can find L such that $E_{N,M,L}$ is of category II. Now for $\vartheta \in E_{N,M,L}$,

$$a \in \mathscr{C}\left(\overline{C}(f, H(\vartheta, -\eta_L))_{+\varepsilon_N/2}\right)$$

and it follows that there is a $\delta = \delta(\vartheta)$ such that the inequality

$$[f(r e^{i\vartheta'}), a] > \varepsilon_N/3 \tag{10}$$

is satisfied for all $\vartheta' \in H(\vartheta, -\eta_L)$, $\vartheta \in E_{N,M,L}$ and $1 - \delta(\vartheta) < r < 1$. We now choose a sequence $\{\delta_p\}$, $\delta_1 > \delta_2 > \cdots > \delta_p > \cdots$, $\lim_{p \to \infty} \delta_p = 0$, and denote by $\mathscr{E}_p$ the sub-set of $E_{N,M,L}$ such that the inequality (10) is satisfied for all $\vartheta' \in H(\vartheta, -\eta_L)$, $\vartheta \in \mathscr{E}_p$ and $1 - \delta_p < r < 1$. Plainly $\mathscr{E}_{p_1} \subseteq \mathscr{E}_{p_2}$ where $p_1 < p_2$ and

$$E_{N,M,L} = \bigcup_p \mathscr{E}_p.$$

It follows that there is a P such that the set $\mathscr{E}_P$ is of category II so that the corresponding points $e^{i\vartheta}$ are dense in an arc γ of the circumference $\varkappa$. Let γ be the open arc $(\vartheta_1, \vartheta_2)$. The inequality (10) is therefore satisfied on the set of radial segments $z = re^{i\vartheta}$,

$$1-\delta_P < r < 1, \quad \vartheta' \in \bigcup_{\vartheta \in \mathscr{E}_P \cap (\vartheta_1, \vartheta_2)} H(\vartheta, -\eta_L);$$

and these radial segments are dense in the annular sector $1-\delta_P < |z| < 1$, $\vartheta_1 < \arg z < \vartheta_2 + \eta_L$ so that, by the continuity of $f(z)$, (10) is satisfied throughout this sector. It follows that $T_M \cap C(f, e^{i\vartheta'})$ *is empty for every* $\vartheta' \in (\vartheta_1, \vartheta_2)$ since every point of T_M is at a distance less than $\varepsilon_N/4$ from a. But since $\mathscr{E}_P$ is dense in $(\vartheta_1, \vartheta_2)$ we may choose $\vartheta' \in \mathscr{E}_P \cap (\vartheta_1, \vartheta_2)$ so that, $\mathscr{E}_P$ being a sub-set of $E_{N,M}$, $T_M \cap C(f, e^{i\vartheta'}) \cap \mathscr{C}\left(C_{H_r}(f, e^{i\vartheta'})_{+\varepsilon_N}\right)$ is not empty and therefore the larger set $T_M \cap C(f, e^{i\vartheta'})$ is not empty. We have thus reached a contradiction which proves the theorem.

Corollary 3. *If $f(z)$ is meromorphic in $|z|<1$ the boundary cluster sets $C_{Br}(f, e^{i\vartheta})$, $C_{Bl}(f, e^{i\vartheta})$ and $C_B(f, e^{i\vartheta})$ are all maximal in a residual set of points $e^{i\vartheta}$ of the circumference $\varkappa$; that is to say $C_{Br}(f, e^{i\vartheta}) = C(f, e^{i\vartheta})$ and similarly for C_{Bl} and C_B.*

8. A property of points of $K(f)$ and $J(f)$

For a given point $e^{i\vartheta} \in K(f)$ we now consider afresh the distribution in the open interval $(-\pi/2, \pi/2)$ of the values of φ for which $C_{\varrho(\varphi)}(f, e^{i\vartheta}) = C_{\mathscr{A}}(f, e^{i\vartheta})$. The same technique enables us to prove

Theorem 8. *If $f(z)$ is meromorphic in $|z|<1$ and if $e^{i\vartheta} \in K(f)$, then $C_{\varrho(\varphi)}(f, e^{i\vartheta}) = C_{\mathscr{A}}(f, e^{i\vartheta})$ for a set of values of φ residual in the open interval $(-\pi/2, \pi/2)$.*

The assertion is of course trivial in the case $e^{i\vartheta} \in F(f)$; and in the case $e^{i\vartheta} \in I(f)$ the set $C_{\mathscr{A}}(f, e^{i\vartheta})$ is total so that the result reduces to a previous theorem[20]).

We shall show that the hypothesis that $C_{\varrho(\varphi)}(f, e^{i\vartheta}) \neq C_{\mathscr{A}}(f, e^{i\vartheta})$ for a set of values of φ of category II in $(-\pi/2, \pi/2)$ implies that $e^{i\vartheta} \in \mathscr{C}K(f)$, which will prove the theorem.

Since $C_{\varrho(\varphi)}(f, e^{i\vartheta})$ is contained in $C_{\mathscr{A}}(f, e^{i\vartheta})$ and is closed the hypothesis $C_{\varrho(\varphi)}(f, e^{i\vartheta}) \neq C_{\mathscr{A}}(f, e^{i\vartheta})$ for a given φ implies that there is a positive ε such that, denoting the closed ε-neighbourhood of $C_{\varrho(\varphi)}(f, e^{i\vartheta})$ by $C_{\varrho(\varphi)}(f, e^{i\vartheta})_{+\varepsilon}$, the set $C_{\mathscr{A}}(f, e^{i\vartheta}) \cap \mathscr{C}\left(C_{\varrho(\varphi)}(f, e^{i\vartheta})_{+\varepsilon}\right)$ is not empty. We again take a decreasing sequence $\{\varepsilon_n\}$ tending to zero and, for a given φ, we denote by $n(\varphi)$ the smallest value of n, if such a number exists, for which $C_{\mathscr{A}}(f, e^{i\vartheta}) \cap \mathscr{C}\left(C_{\varrho(\varphi)}(f, e^{i\vartheta})_{+\varepsilon_n}\right)$ is not empty. For a given $e^{i\vartheta}$ and a given n let B_n be the set of values of φ in $(-\pi/2, \pi/2)$ such that $n(\varphi) = n$. Plainly $B_1 \subseteq B_2 \subseteq \cdots \subseteq B_n \subseteq \cdots$; and if $B = B(\vartheta)$ denotes the set of φ in which $C_{\varrho(\varphi)}(f, e^{i\vartheta}) \neq C_{\mathscr{A}}(f, e^{i\vartheta})$, then

$$B = \bigcup_{1 \leq n < \infty} B_n.$$

[20]) Collingwood (1), p. 181, Theorem 10.

Since by hypothesis B is of category II we can find N such that B_N is of category II in $(-\pi/2, \pi/2)$. We again use the triangular lattice having meshes $T_1, T_2, \ldots, T_m$ all of diameter less than $\varepsilon_N/4$; and for a given $\mu \leq m$ we denote by $B_{N,\mu}$ the subset of B_N in which

$$T_\mu \cap C_{\mathscr{A}}(f, e^{i\theta}) \cap \mathscr{C}\left(C_{\varrho(\varphi)}(f, e^{i\theta})_{+\varepsilon_N}\right)$$

is not empty. Since

$$B_N = \bigcup_{\mu \leq m} B_{N,\mu}$$

we can find M such that $B_{N,M}$ is of category II in $(-\pi/2, \pi/2)$. Choosing $a \in T_M$ we sub-divide $B_{N,M}$ in a way analogous to that employed before. Let $\mathscr{B}_1$ be the sub-set of $B_{N,M}$ such that the inequality

$$[f(e^{i\theta}(1 - te^{i\varphi})), a] > \varepsilon_N/2 \tag{11}$$

is satisfied for all $\varphi \in \mathscr{B}_1$ and $0 < t < \frac{1}{2}$; and, generally, let $\mathscr{B}_p$ be the sub-set of $B_{N,M}$ such that (11) is satisfied for all $\varphi \in \mathscr{B}_p$ and $0 < t < 2^{-p}$. Then

$$B_{N,M} = \bigcup_{1 \leq p < \infty} \mathscr{B}_p$$

and we can find P such that $\mathscr{B}_P$ is of category II in $(-\pi/2, \pi/2)$ and therefore dense in a sub-interval Δ. The angular domain defined by Δ and lying within the circle $|z - e^{i\theta}| < 2^{-P}$ thus contains a dense set of chordal segments on which (11) is satisfied so that it follows from continuity that (11) is satisfied throughout this domain. Hence $C_\Delta(f, e^{i\theta}) \cap T_M$ is empty. But, since $\mathscr{B}_P$ is dense in Δ we may choose $\varphi' \in \mathscr{B}_P \cap \Delta$ so that

$$T_M \cap C_{\mathscr{A}}(f, e^{i\theta}) \cap \mathscr{C}\left(C_{\varrho(\varphi')}(f, e^{i\theta})_{+\varepsilon_N}\right)$$

is not empty and hence also the larger set $T_M \cap C_{\mathscr{A}}(f, e^{i\theta})$ is not empty. Therefore $C_1(f, e^{i\theta}) \neq C_{\mathscr{A}}(f, e^{i\theta})$, so that $e^{i\theta} \in \mathscr{C}K(f)$ and the theorem is proved.

Corollary 4. *If $f(z)$ is meromorphic in $|z| < 1$ and $e^{i\theta} \in J(f)$, then $C_{\varrho(\varphi)}(f, e^{i\theta}) = C(f, e^{i\theta})$ for a set of values of φ residual in $(-\pi/2, \pi/2)$.*

For in $J(f)$ we have $C_{\mathscr{A}}(f, e^{i\theta}) = C(f, e^{i\theta})$ and $J(f) \subseteqq K(f)$.

9. Maximising sub-sets of $J(f)$.

We can choose for the sets $\mathscr{G}_\theta(1), \mathscr{G}_\theta(2), \ldots, \mathscr{G}_\theta(n), \ldots$, of Corollary 1 an enumerable family of curves $\{\lambda_\theta(n)\}$ all having $e^{i\theta}$ as end-point and dense in the circle $|z| < 1$. For example, we may take the family of circular arcs which pass through the points $e^{i\theta}$ and $-e^{i\theta}$. A repetition of the argument of the previous paragraph shows that at every point of $J(f)$ the cluster set on every one of a residual set of these circular arcs is maximal, i.e. equal to $C(f, e^{i\theta})$. By Corollary 3, the sub-set $J_1(f)$ of $J(f)$ in which $C_{B_r}(f, e^{i\theta})$ and $C_{Bl}(f, e^{i\theta})$ are also maximal is again residual in $\varkappa$. If we take $\{\lambda_\theta(n)\}$ to be the family of oricycles through $e^{i\theta}$ the cluster sets upon a residual set of these to the right and left of $e^{i\theta}$ are again maximal in a residual set on $\varkappa$ and

therefore in a residual sub-set of $J_1(f)$; and so on for families of curves in $|z|<1$ having higher orders of contact with the circumference at $e^{i\theta}$. The question naturally arises as to the conditions under which the cluster set on *every curve* to $e^{i\theta}$ is maximal for a significant set of points $e^{i\theta}\in\varkappa$. Simple examples are known for which this is not the case for any point $e^{i\theta}\in\varkappa$. KOENIG's function $\mathscr{K}(z)$ for example, and a function constructed by BAGEMIHL, ERDÖS and SEIDEL are both functions regular in $|z|<1$ for which every point of $\varkappa$ is a limit of zeros of the function and at the same time the end-point of an asymptotic path on which the function tends to infinity. The cluster set of the function is therefore of infinite diameter at every point $e^{i\theta}$ while the cluster set on some curve to this point is degenerate. There are, however, important classes of functions for which the cluster set on every curve to $e^{i\theta}$ is maximal in a residual set of points $e^{i\theta}\in\varkappa$. Apart from these classes the theorems we have proved above are substantially best possible.

We denote by $\Pi(f, e^{i\theta})$ the intersection of the cluster sets $C_\lambda(f, e^{i\theta})$ taken over all possible curves λ in $|z|<1$ ending at $e^{i\theta}$. Then $\Pi(f, e^{i\theta})\subseteqq C(f, e^{i\theta})$; and $\Pi(f, e^{i\theta})$ may be empty as is the case, for example, if $f(z)$ has two distinct asymptotic values at $e^{i\theta}$. We call Π *the principal sub-set of* C and a value $a\in\Pi(f, e^{i\theta})$ *a principal value* at $e^{i\theta}$.

It was proved by GROSS[21]) that for any function $f(z)$ and any point $e^{i\theta}$ a neighbourhood of which is mapped by $f(z)$ on a finite area we have

$$C_{\mathscr{A}}(f, e^{i\theta}) = \Pi(f, e^{i\theta}) \tag{12}$$

so that $\Pi(f, e^{i\theta})$ is certainly not empty at such a point. It follows that for a function $f(z)$ whose Dirichlet integral on $|z|<1$ is finite the equation (12) is satisfied for all $e^{i\theta}$ and the set Π exists at every point of $\varkappa$. Since $C_{\mathscr{A}}(f, e^{i\theta}) = C(f, e^{i\theta})$ for $e^{i\theta}\in J(f)$ we thus have proved

Theorem 9. *If $f(z)$ is meromorphic in $|z|<1$ and if its Dirichlet integral is finite then at every point of the residual set $J(f)$ the cluster set $C(f, e^{i\theta})$ and its principal sub-set are identical.*

Another sufficient condition established by GROSS for the equation (12) is that $C(f, e^{i\theta})$ shall be non-dense so that $C(f, e^{i\theta}) = \mathscr{F}C(f, e^{i\theta})$, where $\mathscr{F}$ denotes the frontier of C. In particular, we have

Theorem 9a. *If $f(z)$ is meromorphic in $|z|<1$ and such that $C(f) = \mathscr{F}C(f)$, then*

$$C(f, e^{i\theta}) = \Pi(f, e^{i\theta})$$

for all $e^{i\theta}\in J(f)$.

In the special case of univalent functions this theorem gives, as a corollary, the complete solution of an old problem of CARATHEODORY on the prime end structure of the boundary of a simply connected domain. This problem is treated elsewhere[22]) and will not therefore be further discussed here.

[21]) GROSS (1), § 8, p. 258. The restriction to regular functions in GROSS' statement of the theorem is superfluous. For $\mathscr{C}R(f, e^{i\theta})$ plainly cannot be empty so that a linear transformation to a function regular in a neighbourhood of $e^{i\theta}$ can always be found.

[22]) See footnote 7.

Part 2

10. Cluster sets on spiral paths

We confine ourselves for convenience to the case of spiral paths in $|z|<1$ converging to the circumference $\varkappa$; but our methods and results are equally valid for more general classes of paths converging to the circumference.

Given any function $f(z)$ meromorphic in $|z|<1$ it is easy to prove that there is an isolated set G of points

$$\{z_n\},\ |z_n|<|z_{n+1}|,\ \lim_{n\to\infty}|z_n|=1,$$

such that $C_G(f)=C(f)$, from which it follows that, given any spiral σ passing through all the points z_n and converging to $\varkappa$, $C_\sigma(f)=C(f)$[23]. This establishes the existence of spirals of maximum indetermination. It can be proved, however, that for certain classes of functions *every* spiral is a path of maximum indetermination.

We recall the following definitions[24]. $R(f)$ is the set of values each taken an infinity of times by $f(z)$ in $|z|<1$, $\mathfrak{C}R(f)$ is its complement and $\overline{\mathfrak{C}R}(f)$ the closure of this set. Similarly, $R(f, e^{i\vartheta})$ is the set of values each of which is taken in every neighbourhood of the point $e^{i\vartheta}$ contained in $|z|<1$, $\Gamma(f)$ is the set of asymptotic values of $f(z)$, and $\Gamma(f)=\Gamma_P(f)\cup\Gamma_A(f)$, where $\Gamma_P(f)$ is the set of values approached on paths having an end point on $\varkappa$ and $\Gamma_A(f)$ the set approached on paths which converge to arcs of the circumference $\varkappa$, of which the spirals are a special case. $\Gamma_P(f, e^{i\vartheta})$ is the set of asymptotic values approached along paths ending in $e^{i\vartheta}$.

We prove the following theorem.

Theorem 10. *If $f(z)$ is meromorphic in $|z|<1$ and such that $\mathfrak{C}R(f)$ contains three points, then, for any spiral path σ in $|z|<1$ converging to $\varkappa$,*

$$C_\sigma(f)\supseteq C(f)\cap\overline{\mathfrak{C}R}(f). \tag{13}$$

Since $\mathfrak{C}R(f)$ contains three points $\Gamma_A(f)$ is empty so that $\Gamma(f)=\Gamma_P(f)$ by a theorem of COLLINGWOOD and CARTWRIGHT[25]. Because σ is a spiral converging to $\varkappa$ and $C_\sigma(f)$ is closed we have $\overline{\Gamma}_P(f)\subseteq C_\sigma(f)$, where $\overline{\Gamma}_P$ denotes the closure of Γ_P. But, by another theorem of COLLINGWOOD and CARTWRIGHT[26], $\overline{\mathfrak{C}R}(f)\cap C(f)\subseteq\overline{\Gamma}(f)$, which proves the theorem.

Corollary 5. *If $f(z)$ is meromorphic in $|z|<1$ and if its Dirichlet integral is finite, then on any spiral σ converging to $\varkappa$,*

$$C_\sigma(f)=C(f). \tag{14}$$

For $\mathfrak{C}R(f)$ must be dense. Otherwise there is a domain δ on the unit w-sphere contained in $R(f)$ and therefore covered an infinity of times by

[23]) COLLINGWOOD (3). I take the opportunity to point out that Theorem 1 of this note had been anticipated by GROSS (1), p. 250, footnote 1.

[24]) COLLINGWOOD and CARTWRIGHT (1).

[25]) loc. cit., p. 97, Theorem 6.

[26]) loc. cit., p. 112, Theorem 9.

the Riemann surface generated by $f(z)$, the area of which is therefore infinite. Hence $\overline{\mathscr{C}R}(f)$ is total and since $C_\sigma(f) \subseteq C(f)$ the result follows.

Corollary 6. *If $f(z)$ is such that $C(f) = \mathscr{F}C(f)$, then*

$$C_\sigma(f) = C(f)$$

for any spiral σ in $|z|<1$ converging to $\varkappa$.

For $\mathscr{C}R(f)$ is dense.

In particular, (14) holds for every function $f(z)$ of finite valency.

Corollary 7. *If $f(z)$ is such that $C(f)$ is sub-total, then for any spiral σ in $|z|<1$ converging to $\varkappa$,*

(15) $$C_\sigma(f) \supseteq \mathscr{F}C(f).$$

For $\mathscr{F}C(f) \subset \overline{\mathscr{C}R}(f)$ and $\mathscr{F}C(f)$ is not empty since $\mathscr{C}C(f)$ is not empty.

This corollary is best possible as is shown by the example of the Blaschke product

$$\Phi(z) = \prod_{n=1}^{\infty} \frac{z - t_n}{1 - t_n z}$$

where

$$t_n = \frac{n! - 1}{n! + 1} \quad (n = 1, 2, \ldots).$$

It was shown by SEIDEL[26a]) that if we consider the sequence of points

$$\tau_n = \frac{(n+1)! - \varrho}{(n+1)! + \varrho}, \quad 1 < \varrho < 2, \; n = 1, 2, \ldots,$$

lying between successive zeros t_n of the function $\Phi(z)$ then

$$\lim_{n \to \infty} \Phi(\tau_n) = \frac{1 - \varrho}{1 + \varrho}.$$

A precisely similar argument shows that for the sequence of points

$$\tau_n' = \frac{(n+1)! - \sqrt{(n+1)}}{(n+1)! + \sqrt{(n+1)}},$$

which also separates successive zeros of $\Phi(z)$, then

$$\lim_{n \to \infty} \Phi(\tau_n') = 1.$$

The point $z = 1$ is the only singularity of $\Phi(z)$ on $\varkappa$ so that $|\Phi(z)| = 1$ at every other point $z \in \varkappa$. It follows that for any $\alpha < 1$ the components of the open set in which $|\Phi(z)| < \alpha$ are infinite in number and are all contained, together with the level curves which are their boundaries, in the open disc $|z| < 1$ and are symmetrical with respect to the real axis. We can therefore trace

[26a]) SEIDEL (1), p. 14.

a spiral σ converging to $\varkappa$ whose only intersections with the positive real axis are points of the sequence $\{\tau_n'\}$ and on which $|\Phi(z)| \to 1$ as $|z| \to 1$. Since the components just referred to are infinite in number $C(\Phi)$ is the closed unit disc Ω so that it follows from Corollary 3 that on this spiral σ, $C_\sigma(\Phi) = \mathfrak{F}\Omega$, which proves our assertion.

There is an alternative proof of Theorem 10 which is of some independent interest. The domain in $|z|<1$ bounded by σ may be mapped conformally onto the disc $|\zeta|<1$ so that the point $\zeta = 1$ corresponds to the circumference $\varkappa$. Denote the mapping function by $\zeta(z)$ and its inverse by $z(\zeta)$. The function $\psi(\zeta) = f(z(\zeta))$ is then meromorphic in $|\zeta|<1$ and satisfies the equations $C(\psi, 1) = C(f)$ and $C_B(\psi, 1) = C_\sigma(f)$.

Suppose that $\mathfrak{C} C_\sigma(f)$ is not empty. By BEURLING's Theorem[26b]), $\mathfrak{F} C(\psi, 1) \subseteqq \mathfrak{F} C_B(\psi, 1)$ from which it follows, since $\mathfrak{C} C_B(\psi, 1) = \mathfrak{C} C_\sigma(f)$ is open, that $\mathfrak{C} C_\sigma(f) \cap \mathfrak{F} C(f)$ is empty. This proves Corollaries 6 and 7 directly.

To prove Theorem 10 we use not BEURLING's Theorem but the well known theorem of IVERSEN and GROSS[27]), by which

$$\mathfrak{C} R(\psi, 1) \cap C(\psi, 1) \cap \mathfrak{C} C_B(\psi, 1) \subseteqq \Gamma_P(\psi, 1). \tag{16}$$

Every path in $|\zeta|<1$ ending at the point $\zeta = 1$ corresponds under the mapping $\zeta(z)$ to a spiral in $|z|<1$ which does not intersect σ, that is to say to a curve in a sub-class of the class of spirals in $|z|<1$. Hence $\Gamma_P(\psi, 1) \subseteqq \Gamma_S(f)$, where $\Gamma_S(f)$ denotes the sub-set of asymptotic values of $f(z)$ which are approached along spirals. From (16) we have therefore

$$\mathfrak{C} R(f) \cap C(f) \cap \mathfrak{C} C_\sigma(f) \subseteqq \Gamma_S(f). \tag{17}$$

Now, by a theorem of COLLINGWOOD and CARTWRIGHT quoted above[28]) $\Gamma_S(f)$ is empty if $\mathfrak{C} R(f)$ contains three values and so under this condition

$$\mathfrak{C} R(f) \cap C(f) \subseteqq C_\sigma(f) \tag{18}$$

and hence, since $C(f)$ and $C_\sigma(f)$ are both closed,

$$\overline{\mathfrak{C} R}(f) \cap C(f) \subseteqq C_\sigma(f),$$

which proves the theorem.

We should note that if $\mathfrak{C} R(f)$ contains only two points both of them may belong to $\Gamma_S(f)$[29]) so that in this case we can find spirals σ_1 and σ_2 such that $C_{\sigma_1}(f) \cap C_{\sigma_2}(f)$ is empty.

A theorem similar to Theorem 10 may be proved by our first method for a path which converges to an arc of $\varkappa$.

[26b]) See COLLINGWOOD and CARTWRIGHT (1), p. 123.
[27]) See COLLINGWOOD and CARTWRIGHT (1), p. 124, Theorem C'.
[28]) loc. cit., p. 97, Theorem 6.
[29]) See VALIRON (1) for an example.

11. A method of construction

The mapping of the spiral domain described in the previous paragraph gives a method for the construction of a function whose cluster set at the point $\zeta = 1$ can be prescribed within the limitations imposed by Theorem 10. The most important of these lies in the fact that, given an arbitrary continuum C, it is not true in general that there is a meromorphic function $f(z)$ such that $C(f) = C$. A known counter-example of such a continuum due to POTYAGAILO and RUDIN is in fact non-dense[30]). However, the frontier of any simply connected domain, however complicated its structure, is an admissible continuum and may be transformed from $C(f)$ to $C(\psi, 1)$ by some standard mapping, since the spiral in this case is arbitrary.

As we have seen, given any function $f(z)$ meromorphic in $|z| < 1$, there is some spiral σ such that $C_\sigma(f) = C(f)$; but since the form of σ depends upon $f(z)$ there can be no standard mapping for the actual construction of $\psi(\zeta)$ but a mere existence theorem in the general case.

12. Generalisations

A domain of finite connectivity $p > 1$ may be mapped conformally upon a domain of the same connectivity belonging to any one of a variety of canonical types, parallel slit domains, radial slit domains, circle domains, etc.[31]). We choose the circle domain, that is to say a domain containing the point at infinity whose complement consists of disjoint circular discs, as the canonical type to consider. We shall consider domains having only non-degenerate boundary components since the properties of an analytic function at an isolated boundary point are fully known. The domains which may be conformally mapped upon such a canonical domain have no degenerate boundary component.

Let D be the circle domain and denote the circular boundary components by $\varkappa_1, \varkappa_2, \ldots, \varkappa_p$, so that $\mathfrak{F}D = \bigcup_{\nu \leq p} \varkappa_\nu$. Suppose that $f(z)$ is meromorphic in D. The cluster sets of $f(z)$ on a component $\varkappa_\nu$, denoted by $C(f, \varkappa_\nu)$, and at a point $\zeta \in \varkappa_\nu$ denoted by $C(f, \zeta)$, are defined in the obvious way, e.g. $a \in C(f, \varkappa_\nu)$ if there is a sequence $\{z_n\}$, $z_n \in D$, $\lim_{n\to\infty} d(z_n, \varkappa_\nu) = 0$, such that $\lim_{n\to\infty} f(z_n) = a$, where $d(z, \varkappa)$ is the distance from the point z to the set $\varkappa$. Now $\varkappa_\nu$ may be enclosed in a concentric circle χ_ν which does not meet $\mathfrak{F}D$ or contain any other component of $\mathfrak{F}D$, so that $f(z)$ is meromorphic in the annulus between $\varkappa_\nu$ and χ_ν. Cut this annulus along a radius and map the resulting simply connected domain on the circular disc $|\xi| < 1$. $\varkappa_\nu$ is then mapped upon an arc γ_ν of the circle $|\xi| = 1$ and the mapping is conformal on the two arcs $\varkappa_\nu$ and γ_ν themselves. All the theorems of §§ 4–9 above, together with the

[30]) POTYAGAILO (1); RUDIN (1).
[31]) See for example JULIA (1).

theorems of FATOU and PLESSNER, are valid for the function $\psi(\xi)=f(z(\xi))$ in $|\xi|<1$, where $z(\xi)$ is the inverse of the function $\xi(z)$ which maps the annulus on $|\xi|<1$. Since the mapping is conformal on $\varkappa_\nu$ and γ_ν it follows that to the sets $J(\psi)\cap\gamma_\nu$ and $K(\psi)\cap\gamma_\nu$ there correspond sets on $\varkappa_\nu$ which we may call $J(f,\varkappa_\nu)$ and $K(f,\varkappa_\nu)$ such that, if Δ denotes an angular domain in D having its vertex at $\zeta\in\varkappa_\nu$, we have $C_\Delta(f,\zeta)=C(f,\zeta)$ for $\zeta\in J(f,\varkappa_\nu)$ and, given any two angular domains Δ_1 and Δ_2 having the common vertex ζ, we have $C_{\Delta_1}(f,\zeta)=C_{\Delta_2}(f,\zeta)$ for $\zeta\in K(f,\varkappa_\nu)$. At corresponding points $\zeta\in\varkappa_\nu$ and ξ the equation $C(f,\xi)=C(\psi,\xi)$ is satisfied and for $\zeta\in K(f,\varkappa_\nu)$ and its corresponding point $\xi(\zeta)\in K(\psi)\cap\gamma_\nu$ every angular cluster set $C_\Delta(f,\zeta)$ is equal to every angular cluster set $C_\Delta(\psi,\xi)$ in virtue of the definitions of $K(\psi)$ and $K(f,\varkappa_\nu)$. We denote by $F(f,\varkappa_\nu)$ the sub-set of $K(f,\varkappa_\nu)$ in which the angular cluster sets $C_\Delta(f,\zeta)$ are degenerate and by $I(f,\varkappa_\nu)$ the sub-set in which they are total. Since sets of zero measure and sets of category I on $\varkappa_\nu$ correspond respectively to sets of zero measure and sets of category I on γ_ν an application of PLESSNER's theorem and of Theorem 4 to $\psi(\xi)$ proves the following theorem.

Theorem 11. *If $f(z)$ is meromorphic in D and $\varkappa_\nu$ is one of its circular boundaries, then the set $K(f,\varkappa_\nu)$ is residual on $\varkappa_\nu$ and the sub-set $F(f,\varkappa_\nu)\cup I(f,\varkappa_\nu)$ of $K(f,\varkappa_\nu)$ is* p.p. *on $\varkappa_\nu$.*

An alternative method is to prove an analogue of Theorem 1 for a contour $\varkappa_\nu$ and sets $\mathcal{G}_\vartheta$ in D, where now ϑ is a rotation about the centre of $\varkappa_\nu$. The proof is the same as that of Theorem 1 apart from trivial verbal changes. We can then proceed again to handle every individual contour by mapping the cut annulus between $\varkappa_\nu$ and χ_ν on $|\xi|<1$ and thus prove for each $\varkappa_\nu$ the analogues of all the theorems on pointwise cluster sets proved in Part 1 which are not implied by Theorem 11. Our method is not necessarily applicable to a domain D of infinite connectivity for then $\varkappa_\nu$ may not be an isolated boundary component. The properties of $f(z)$ near a non-isolated boundary component depend upon its relation to the neighbouring boundary components[32]).

Some of the results may be usefully extended to more general domains whose boundary components are not circular. Consider a domain $\mathcal{D}$ of finite connectivity whose complement contains p non-degenerate components whose frontiers are denoted by $k_1, k_2, \ldots, k_p$ and which may accordingly be mapped conformally upon a domain D having p corresponding circular boundary components $\varkappa_1, \varkappa_2, \ldots, \varkappa_p$. There is a 1:1 correspondence under the mapping between points of $\varkappa_\nu$ and frontier elements (prime ends) of $\mathcal{D}$ in k_ν and it follows from our theorems that on any given k_ν there is a residual set of frontier elements which are the "ends" of a set of paths $\{\mu(\vartheta,k_\nu)\}$ corresponding under the mapping to a set $\{\lambda(\vartheta,\varkappa_\nu)\}$ of paths of fixed form,

[32]) This is illustrated by the conformal mapping of domains of infinite connectivity upon other plane domains, which has been studied by a number of authors. For recent results and references to previous work see STREBEL (1) and BROWN (1).

e.g. radial paths in D terminating on $\varkappa_\nu$ in a residual set, and on which the cluster set $C_{\mu(\vartheta, k_\nu)}(\psi, \omega)$ is maximal, i.e. $C_{\mu(\vartheta, k_\nu)}(\psi, \omega) = C(\psi, \omega)$, where $\omega \in k_\nu$ is a frontier element not necessarily a point. The set $\{\mu(\vartheta, k_\nu)\}$ intersects every closed curve obtained by the mapping of a circle χ_ν sufficiently close to $\varkappa_\nu$ in a residual set and, adapting a term of BAGEMIHL and SEIDEL[33]), may be called a residual tress. If we consider the mapping function $\zeta(z)$ which maps D on $\mathscr{D}$ our method at once gives for each k_ν the same theorems on the distribution of the four classes of prime ends as we have proved for a simply connected domain.

13. Turning now to Part 2, we repeat the mapping of the cut annulus between $\varkappa_\nu$ and χ_ν onto the unit circle $|\xi| < 1$ and we observe that the theorem of GROSS and IVERSEN applies to the function $\psi(\xi) = f(z(\xi))$ on $|\xi| = 1$. The mapping puts a spiral σ_ν in D whose closure contains $\varkappa_\nu$ onto a sequence Σ of curves in $|\xi| < 1$ converging to the arc γ_ν defined above and separating γ_ν from the origin $\xi = 0$. Plainly the cluster set $C_\Sigma(\psi)$ of $\psi(\xi)$ on Σ is equal to $C_{\sigma_\nu}(f, \varkappa_\nu)$ and we can apply to $\psi(\xi)$ on Σ the proof of Theorem 10 with only verbal changes and thus prove

Theorem 12. *If $f(z)$ is meromorphic in a domain D of finite connectivity and if, for a non-degenerate circular boundary component $\varkappa_\nu$, $\mathscr{C}R(f, \varkappa_\nu)$ contains three points, then for any spiral σ_ν in D converging to $\varkappa_\nu$, we have*

$$C_{\sigma_\nu}(f) \geqq C(f, \varkappa_\nu) \cap \overline{\mathscr{C}R}(f, \varkappa_\nu). \tag{19}$$

This leads, with the obvious verbal adjustments, to a group of corollaries exactly parallel to Corollaries 5, 6 and 7 given in § 10. In contrast with Theorem 12 and its corollaries we may recall a theorem of BAGEMIHL and SEIDEL[34]) asserting the existence of a function meromorphic in D and having a prescribed enumerable set of asymptotic values on a prescribed set of spirals in D converging to one of the boundary components $\varkappa_\nu$.

14. It will have been noted that the only property of analytic functions used in the proof of Theorem 1 is the property of being a continuous open mapping. On the other hand, FATOU's theorem and therefore also PLESSNER's theorem depend upon the Poisson integral representation of a harmonic function. The theorem of GROSS and IVERSEN also requires more than continuity. But all our theorems on maximal cluster sets which do not depend on the FATOU-PLESSNER or GROSS-IVERSEN theorems can be immediately extended to continuous functions. In particular, Theorems 3 and 4 asserting that $J(f)$ and $K(f)$ are residual sets on $\varkappa$ are valid for continuous functions while FATOU's and PLESSNER's theorems require further conditions. Theorem 4, for example, is in this sense more general than PLESSNER's theorem.

[33]) BAGEMIHL and SEIDEL (3).
[34]) BAGEMIHL and SEIDEL (2).

References

BAGEMIHL, F., and W. SEIDEL: (1) A general principle involving Baire category, with applications to function theory and other fields. Proc. Nat. Acad. Sci. U.S.A. **39**, 1068—1075 (1953). — (2) Spiral and other asymptotic paths, and paths of complete indetermination, of analytic and meromorphic functions. Proc. Nat. Acad. Sci. U.S.A. **39**, 1251—1258 (1953). — (3) Some boundary properties of analytic functions. Math. Z. **61**, 186—199 (1954). — BROWN, L. J. M.: (1) On conformal mappings of domains of infinite connectivity. Proc. Cambridge Phil. Soc. **51**, 56—64 (1955). — COLLINGWOOD, E. F.: (1) On the linear and angular cluster sets of functions meromorphic in the unit circle. Acta math. **91**, 165—185 (1954). — (2) Sur le comportement à la frontière, d'une fonction méromorphe dans le cercle unité. C. R. Acad. Sci. Paris **240**, 1502—1504 (1955). — (3) Sur les ensembles d'indétermination maximum des fonctions analytiques. C. R. Acad. Sci. Paris **240**, 1604—1606 (1955). — (4) A theorem on certain classes of singularities defined by cluster sets. J. London Math. Soc. **30**, 422—424 (1955). — (5) On a theorem of Eggleston concerning cluster sets. J. London Math. Soc. **30**, 425—428 (1955). — (6) A theorem on prime ends. J. London Math. Soc. **31**, 344—349 (1956). — COLLINGWOOD, E. F., and M. L. CARTWRIGHT: (1) Boundary theorems for a function meromorphic in the unit circle. Acta math. **87**, 83—146 (1952). — EGGLESTON, H. G.: (1) A property of bounded analytic functions. Comment. Math. Helv. **30**, 139—143 (1956). — GROSS, W.: (1) Über die Singularitäten analytischer Funktionen. Monatsb. Math. u. Phys. **29**, 3—47 (1918). — (2) Zum Verhalten analytischer Funktionen in der Umgebung singulärer Stellen. Math. Z. **2**, 243—294 (1918). — JULIA, G.: (1) Leçons sur la représentation conforme des aires multiplement connexes. Paris 1934. — PLESSNER, A.: (1) Über das Verhalten analytischer Funktionen am Rande ihres Definitionsbereiches. J. reine u. angew. Math. **158**, 219—227 (1927). — POTYAGAILO, D. B.: (1) On the set of boundary values of meromorphic functions. Doklady Acad. Nauk SSSR., N. S. **86**, 661—663 (1952) [Russian]. — RUDIN, W.: (1) On a problem of COLLINGWOOD and CARTWRIGHT. J. London Math. Soc. **30**, 231—238 (1955). — SEIDEL, W.: (1) On the cluster values of analytic functions. Trans. Amer. Math. Soc. **34**, 1—21 (1932). — STOILOW, S.: (1) Leçons sur les principes topologiques de la théorie des fonctions analytiques. Paris 1938. — STREBEL, K.: (1) Über die konforme Abbildung von Gebieten unendlich hohen Zusammenhangs. Comm. Math. Helv. **27**, 101—127 (1953). — VALIRON, G.: (1) Sur les singularités des fonctions holomorphes dans un cercle. C. R. Acad. Sci. Paris **198**, 2065—2067 (1934). — WHYBURN, G. T.: (1) Open mappings on locally compact spaces. Memoirs Amer. Math. Soc. **1950**, No 1, 23 pp.

Lilburn Tower, Alnwick, Northumberland, England

(Eingegangen am 15. Juni 1956)

BRAUER, G.
Math. Zeitschr. Bd. 67, S. 397—403 (1957)

Some Summation Matrices of Hausdorff Type*)

By

GEORGE BRAUER

1. Introductory remarks. The purpose of this paper is to study certain methods of assigning generalized limits to divergent sequences of complex numbers. All methods considered here are represented by infinite triangular matrices $A = (a_{n,k})$. The A-transform $A\{s_n\}$ of the sequence $\{s_n\}$ is the sequence t_n where

$$t_n = \sum_{k=0}^{n} a_{n,k} s_k .$$

If the sequence t_n converges to some complex number σ, then A is said to evaluate $\{s_n\}$ to σ, and $\{s_n\}$ is said to lie in the convergence field of A. If the matrix A evaluates every convergent sequence to its limit, it is said to be regular. TOEPLITZ [6] proved that for a matrix $A = (a_{n,k})$ to be regular it is necessary and sufficient that

(i) $$\lim_{n\to\infty} a_{n,k} = 0 \qquad (k = 0, 1, \ldots),$$

(ii) $$\sup \sum_{k=0}^{n} |a_{n,k}| < \infty ,$$

(iii) $$\lim_{n\to\infty} \sum_{k=0}^{n} a_{n,k} = 1 .$$

The quantity $\limsup \sum_{k=0}^{n} |a_{n,k}|$ will be called the norm of A and will be written $\|A\|$. In order that a triangular matrix transform every bounded sequence into a bounded sequence it is necessary and sufficient that $\|A\| < \infty$.

An interesting class of summation matrices is the class of Hausdorff matrices, that is, the class of matrices of the form $\varrho\mu\varrho^{-1}$, where μ is an arbitrary diagonal matrix and ϱ is the self-reciprocal Euler matrix with elements $(\varrho_{n,k})$ given by the relations

$$\varrho_{n,k} = (-1)^k \binom{n}{k} .$$

The sequence $\{\mu_n\}$ appearing on the main diagonal of μ shall be said to generate the Hausdorff matrix $\varrho\mu\varrho^{-1}$.

*) Submitted as a dissertation in partial fulfillment of the requirements for the degree of Doctor of Philosophy at the University of Michigan. The author wishes to thank Professor GEORGE PIRANIAN for many helpful suggestions.

The theory of Hausdorff matrices is developed in [2]. It is shown there that if $\{\mu_n\}$ generates a Hausdorff matrix of finite norm, then there exists a function, $t(u)$, of bounded variation on $0 \leq u \leq 1$ which satisfies the equations

$$\mu_n = \int_0^1 u^n \, dt(u), \qquad n = 0, 1, \ldots .$$

In terms of $t(u)$, the elements of $\varrho^{-1}\mu\varrho$ are given by the equations

$$a_{n,k} = \binom{n}{k} \int_0^1 u^k (1-u)^{n-k} \, dt(u).$$

The best-known regular Hausdorff matrix is the Cesàro matrix. It is generated by the sequence $\{1/(n+1)\}$, and its elements $c_{n,k}$ are given by the relations

$$c_{n,k} = 1/(n+1), \qquad k \leq n, \; n = 0, 1, \ldots,$$

$$c_{n,k} = 0, \qquad k > n.$$

In [3] Hausdorff extends his theory to some matrices of the form $A\mu A^{-1}$ where μ is a diagonal matrix and A is a normal matrix. I shall call a matrix of this form a matrix of HAUSDORFF; the sequence appearing on the main diagonal of μ will always be denoted by $\{\mu_n\}$; this sequence will be said to generate the matrix $A\mu A^{-1}$.

In this paper I shall study certain classes of matrices of Hausdorff type. Unlike many regular Hausdorff matrices, the regular matrices of Hausdorff type considered here have very small convergence fields, and, in most cases, they evaluate no divergent sequences. I do not know the full reason for the situation. I believe that it is partly due to the fact that for most matrices $A\mu A^{-1}$ studied here, the matrix A is regular, unlike the Euler matrix.

We shall use the following simple facts about matrices of Hausdorff type throughout the paper.

(a) If μ is a diagonal matrix, A is a normal matrix, and $\|A\mu A^{-1}\| < \infty$ then $\{\mu_n\}$ is a bounded sequence.

(b) If μ is a diagonal matrix, A is a regular normal matrix, and $A\mu A^{-1}$ is regular, then $A\{\mu_n\} \to 1$.

(c) If μ is a diagonal matrix, A is a normal matrix whose reciprocal is regular, and $A\mu A^{-1}$ is regular, then $\mu_n \to 1$.

To prove (a) we let $A = (a_{n,k})$ $A^{-1} = (\alpha_{n,k})$, and $A\mu A^{-1} = (d_{n,k})$. Then $d_{n,n} = a_{n,n}\,\mu_n\,\alpha_{n,n} = \mu_n$ and the result follows.

To prove (b) we note that since A and $A\mu A^{-1}$ are regular, the matrix $(A\mu A^{-1})A = A\mu$ is regular. Consequently, $\sum_{k=0}^{n} a_{n,k}\,\mu_k \to 1$ as $n \to \infty$, that is, $A\{\mu_n\} \to 1$.

To prove (c) we observe that our assumptions imply that $\mu A^{-1}=A^{-1}$ $(A\mu A^{-1})$ is regular. Since $\mu A^{-1}=(\mu_n\alpha_{n,k})$ we have $\lim_{n\to\infty}\mu_n\Big(\sum_{k=0}^{n}\alpha_{n,k}\Big)=1$. Since $\lim_{n\to\infty}\sum_{k=0}^{n}\alpha_{n,k}=1$, $\mu_n\to 1$.

1. *The Matrices* $E_\alpha\mu E_\alpha^{-1}$: Let α denote any complex number other than 0. Then the Euler Knopp matrix of order α, which we denote by E_α, is defined as the Hausdorff matrix generated by the sequence $\{\alpha^n\}$; its elements are given by the formula

$$e_{n,k}=\binom{n}{k}\alpha^k(1-\alpha)^{n-k}$$

cf. ([5] pp. 245).

THEOREM 2.1. *The elements* $b_{n,k}$ *of the matrix* $E_\alpha\mu E_\alpha^{-1}$ *are given by the formula*

$$b_{n,k}=(1-\alpha)^{n-k}\binom{n}{k}\Delta^{n-k}\mu_k. \tag{1}$$

Proof. Let $E_\alpha^{-1}=(\vartheta_{n,k})$. We then have

$$\vartheta_{n,k}=\binom{n}{k}(\alpha^{n-k}-1)/\alpha^n$$

and thus

$$\begin{aligned}
b_{n,k}&=\sum_{j=k}^{n}e_{n,j}\mu_j\vartheta_{j,k}\\
&=\sum_{j=k}^{n}\binom{n}{j}\alpha^j(1-\alpha)^{n-j}\mu_j\binom{j}{k}(\alpha-1)^{j-k}/\alpha^j\\
&=\binom{n}{k}(1-\alpha)^{n-k}\sum_{j=k}^{n}(-1)^{j-k}\binom{n-k}{n-j}\mu_j\\
&=(1-\alpha)^{n-k}\binom{n}{k}\Delta^{n-k}\mu_k.
\end{aligned}$$

If $\alpha=0$, then the matrix $E_\alpha\mu E_\alpha^{-1}$ is undefined. However if α is set equal to 0 in (1), then the resulting expression yields the elements of the Hausdorff matrix generated by $\{\mu_n\}$.

It is clear that if $\|E_\alpha\mu E_\alpha^{-1}\|<\infty$, and β is a complex number such that $|1-\beta|\leq|1-\alpha|$, then $\|E_\beta\mu E_\beta^{-1}\|<\infty$. In particular, if $\{\mu_n\}$ generates a Hausdorff matrix of finite norm, and $|1-\alpha|\leq 1$, then $\|E_\alpha\mu E_\alpha^{-1}\|<\infty$.

The following result is an immediate consequence of Theorem 2.1.

THEOREM 2.2. *If* $\|E_\alpha\mu E_\alpha^{-1}\|<\infty$, *then*

$$\Delta^r\mu_n=O(n^{-r}),\qquad r=0,1,\ldots.$$

THEOREM 2.3. *If* $\alpha>0$ *and* $E_\alpha\mu E_\alpha^{-1}=(b_{n,k})$ *is regular, then* $\mu_n\to 1$.

Proof. If $\alpha\geq 1$, then E_α^{-1} is regular and it follows from (c) on p. 398 that $\mu_n\to 1$. If $\alpha<1$, then E_α is regular. By (b) on p. 398, $E_\alpha\{\mu_n\}\to 1$. Since $\Delta\mu_n=O(1/n)$ by Theorem 2.2 it follows that $\mu_n\to 1$ (cf. [5], pp. 139—141).

THEOREM 2.4. *If* $\alpha>0$ *and* $E_\alpha \mu E_\alpha^{-1}=(b_{n,k})$ *is regular, then*

$$\lim_{n\to\infty} b_{n,n-r}=0, \qquad r=1,2,\ldots.$$

Proof. It is sufficient to show that $\Delta^r \mu_n = o(n^{-r})$ for all positive integers r. To accomplish this we require two lemmas.

LEMMA A. *If* $\alpha \geq 0$ *and* $\{u_n\}$ *is a sequence such that*

(i) $$u_n = o(n^{-\alpha}),$$

(ii) $$\Delta^2 u_n = O(n^{-\alpha-2}),$$

then

$$\Delta u_n = o(n^{-\alpha-1}).$$

Proof. We have

$$u_{n+h} = u_n - h\,\Delta u_n + \sum_{j=0}^{h-2} (h-j-1)\,\Delta^2 u_{n+j}$$

so that

(1) $$|\Delta u_n| \leq (1/h)\left[|u_{n+h}-u_n| + h^2 \max_{0\leq j\leq h} |\Delta^2 u_{n+j}|\right].$$

Let ε be an arbitrary positive number and let N be chosen so large that

$$|u_N - u_n| < \varepsilon\, n^{-\alpha} \qquad \text{for} \quad n \geq N.$$

If we set $h=[n\sqrt{\varepsilon}]$, then, by (1) we have[1])

$$|\Delta u_n| < (M+1)\sqrt{\varepsilon}\, n^{-\alpha-1}, \qquad \text{where} \qquad M = \sup\left|n^{\alpha+2}\Delta^2 u_n\right|.$$

Since ε is arbitrary, we may conclude that $\Delta u_n = o(n^{-\alpha-1})$.

LEMMA B. *If* v_n *is a null sequence such that*

$$\Delta^r v_n = O(n^{-r}) \qquad r=1,2,\ldots,$$

then

$$\Delta^r v_n = o(n^{-r}) \qquad r=1,2,\ldots.$$

Proof. We use mathematical induction. To show that $\Delta v_n = o(1/n)$ we use Lemma A with $\alpha=0$, $u_n=v_n$. Let us suppose that $\Delta^{r_0} v_n = o(n^{-r_0})$. If we apply Lemma A with $\alpha=r_0$, $u_n=\Delta^{r_0}v_n$, we see that $\Delta^{r_0+1}v_n = \Delta(\Delta^{r_0}v_n) = o(n^{-r_0-1})$. Hence the induction is complete.

To complete the proof of Theorem 2.4 we need only apply Lemma B with $u_n=\mu_n-1$.

THEOREM 2.5. *If* $\alpha>0$, $E_\alpha \mu E_\alpha^{-1}$ *is a regular matrix, and* β *is a complex number such that* $|1-\beta|<|1-\alpha|$, *then* $E_\beta \mu E_\beta^{-1}$ *is a regular matrix which evaluates no divergent sequences. Moreover* $E_\beta \mu E_\beta^{-1}$ *transforms every bounded sequence* $\{s_n\}$ *into a sequence which differs from* $\{s_n\}$ *by a null sequence.*

Proof. Let $E_\alpha \mu E_\alpha^{-1}=B=(b_{n,k})$ and let $E_\beta \mu E_\beta^{-1}=B'=(b'_{n,k})$. Then $b'_{n,k}=\left[(1-\beta)/(1-\alpha)\right]^{n-k} b_{n,k}$. Let ε be an arbitrary positive number, and let

[1]) $[x]$ denotes the greatest integer less than x.

R be a positive integer such that $[(1-\beta)/(1-\alpha)]^R<\varepsilon$. Now let N be chosen so that for $n\geq N$, $k>n-R$

$$|b_{n,k}|\leq \varepsilon/R,$$

and

$$\sum_{k=0}^{n}|b_{n,k}|\leq \|B\|+1.$$

For $n\geq N$ we have

$$\begin{aligned}\sum_{k=0}^{n-1}|b'_{n,k}| &\leq \sum_{k=0}^{n-1}\left|\left(\frac{1-\beta}{1-\alpha}\right)^{n-k}b_{n,k}\right| \\ &\leq \left|\frac{1-\beta}{1-\alpha}\right|^R\sum_{k=0}^{n-R}|b_{n,k}|+\sum_{k=n-R+1}^{n-1}|b_{n,k}| \\ &\leq (\|B\|+1)\,\varepsilon+R\cdot\varepsilon/R=(\|B\|+2)\,\varepsilon.\end{aligned}$$

Since ε is an arbitrary positive number,

$$\lim_{n\to\infty}\sum_{k=0}^{n-1}|b'_{n,k}|=0.$$

Also

$$\lim_{n\to\infty}b'_{n,n}=\lim_{n\to\infty}\mu_n=1.$$

Therefore B' is regular, and it follows from a theorem of AGNEW ([1], pp. 245 to 246), that B' evaluates no divergent sequences.

Let $\{s_n\}$ be a bounded sequence and let $t_n=B'\{s_n\}$. Then

$$t_n=\sum_{k=0}^{n-1}b'_{n,k}s_k+b'_{n,n}s_n.$$

Since $\{s_n\}$ is a bounded sequence, $\sum_{k=0}^{n-1}b'_{n,k}s_k$ tends to 0. Also, since $b_{n,n}$ tends to 1, $\{s_n-t_n\}$ is a null sequence.

It remains an open question whether there exists a positive number α and a sequence $\{\mu_n\}$ such that the matrix $E_\alpha\mu E_\alpha^{-1}$ is regular and evaluates a divergent sequence.

THEOREM 2.6. *If* $|1-\alpha|>1$ *and* $E_\alpha\mu E_\alpha^{-1}$ *is regular, then it is the identity matrix.*

Proof. Let ϱ denote the Euler matrix and let $\varrho\mu\varrho^{-1}=A=(a_{n,k})$. Then

$$a_{n,k}=\binom{n}{k}\Delta^{n-k}\mu_k.$$

If we let $E_\alpha\mu E_\alpha^{-1}=B=(b_{n,k})$

$$a_{n,k}=\frac{b_{n,k}}{(1-\alpha)^{n-k}},$$

and hence

$$|a_{n,k}|<|b_{n,k}|.$$

Since $\|B\| < \infty$, A is a Hausdorff matrix of finite norm. Consequently, there exists a function $t(u)$, of bounded variation in the interval $0 \leq u \leq 1$, such that

$$a_{n,k} = \binom{n}{k} \int_0^1 u^k (1-u)^{n-k} \, dt(u)$$

and, for each number p such that $0 < p < 1$

$$\lim_{n \to \infty} \sum_{0 \leq k \leq pn} |a_{nk}| = \int_0^p |dt(u)|$$

cf. ([2], pp. 104—109).

Let p denote a fixed number in $[0, 1)$ and let ε be an arbitrary positive number. Then there exists an integer N such that if $n \geq N$, then

$$\sum_{0 \leq k \leq pn} |a_{n,k}| \geq \int_0^p |dt(u)| - \varepsilon.$$

Hence

$$\sum_{0 \leq k \leq pn} |b_{n,k}| \geq |1-\alpha|^{n(1-p)} \left[\int_0^p |dt(u)| - \varepsilon \right].$$

Since $\sum_{k=0}^{n} |b_{n,k}|$ is bounded and the first member on the right tends to infinity the second factor cannot be positive for any value of p. Since ε is arbitrary, $\int_0^p |dt(u)| = 0$ for $0 \leq p < 1$. Hence $t(u)$ is constant except for a possible saltus at $t = 1$. The result follows from the formula for $b_{n,k}$.

3. The Matrices $C_m \mu C_m^{-1}$. For each nonnegative integer m the matrix $C_m \mu C_m^{-1}$ is defined as the Hausdorff matrix generated by the sequence $\left\{1 / \binom{n+m}{n}\right\}$ (cf. [2], pp. 82—83). In this section I shall list the main results on matrices $C_m \mu C_m^{-1}$, where m always denotes a nonnegative integer. For brevity we shall denote $C_m \mu C_m^{-1}$ by $Q^{(m)}$ and its elements by $(q_{n,k}^{(m)})$.

THEOREM 3.1. *For each non-negative integer m, the elements of $Q^{(m)}$ are given by*

$$(1) \qquad q_{n,k}^{(m)} = \left(\prod_{i=1}^{m} \frac{k+i}{n+i} \right) \sum_{j=k}^{k+m} (-1)^{j-k} \binom{m}{j-k} \binom{n-j+m-1}{n-j} \mu_j,$$

$$(2) \qquad q_{n,k}^{(m)} = \left(\prod_{i=1}^{m} \frac{k+i}{n+i} \right) \Delta^m \left[\binom{n-k+m-1}{n-k} \mu_k \right],$$

$$(3) \qquad q_{n,k}^{(m)} = \left(\prod_{i=1}^{m} \frac{k+i}{n+i} \right) \sum_{r=0}^{m-1} \binom{m}{r} \binom{n-k+m-r-1}{n-k} \Delta \mu_{k+r}^{m-r}.$$

The following relation exists between the elements of $C_m \mu C_m^{-1}$ and those of $C_{m-1} \mu C_{m-1}^{-1}$.

$$(4) \qquad q_{n,k}^{(m)} = \frac{(k+m)(n-k+m-1) q_{n,k}^{(m-1)} - (k+1)(n-k-1) q_{n,k+1}^{(m-1)}}{(m-1)(n+m)}$$

and this formula enables us to prove

THEOREM 3.2. *If* $m \geq 1$, *and* $\| C_m \mu C_m^{-1} \| < \infty$, *then* $\| C_{m-1} \mu C_{m-1}^{-1} \| < \infty$.

The relation (4) also helps us to obtain the following result:

THEOREM 3.3. *If* $m > 1$ *and* $\| Q^{(m)} \| < \infty$ *then*

$$\Delta^p \mu_n = O(n^{-p}), \qquad p = 0, 1, \ldots, m-1.$$

If $m > 1$ and $Q^{(m)}$ is regular then $C_m\{\mu_n\} \to 1$. By the above result $\Delta \mu_n = O(1/n)$. Using a well known Tauberian theorem we obtain

THEOREM 3.4. *If* $m > 1$ *and* $Q^{(m)}$ *is regular, then* $\mu_n \to 1$.

Using Theorems 3.3 and 3.4 together with (3) applied to $(Q^{(m)})^{-1}$ we obtain

THEOREM 3.5. *If* $m > 1$ *and* $Q^{(m)}$ *is regular, then it evaluates no devergent sequences. A regular matrix* $C_1 \mu C_1^{-1}$ *evaluates divergent sequences if and only if* 0 *is a cluster value of* $\{\mu_n\}$.

References

[1] AGNEW, R. P.: On Equivalence of Methods of Evaluation of Sequences. Tôhoku Math. J. **35**, 244—252 (1935). — [2] HAUSDORFF, FELIX: Summationsmethoden und Momentfolgen. I. Math. Z. **9**, 74—109 (1921). — [3] HAUSDORFF, FELIX: Summationsmethoden und Momentfolgen. II. Math. Z. **9**, 280—299 (1921). — [4] KNOPP, KONRAD: Über das Eulersche Summierungsverfahren. I. Math. Z. **15**, 226—253 (1922). — [5] KNOPP, KONRAD: Über das Eulersche Summierungsverfahren. II. Math. Z. **18**, 125—156 (1923). — [6] TOEPLITZ, OTTO: Über allgemeine lineare Mittelbildungen. Prace Mat-Fiz. **22**, 113—119 (1911).

University of Minnesota, Minneapolis, Minnesota, USA

(Eingegangen am 28. Mai 1956)

SRIVASTAVA, K. J.
Math. Zeitschr. Bd. 67, S. 404—412 (1957)

Fractional integration and Meijer transform

By

KRISHNA JI SRIVASTAVA

1. In 1940 ERDÉLYI and KOBER [*1–2* and *3–4*] used certain operators of fractional integration and differentiation to develope the theory of the well-known Hankel transform of functions belonging to the class $L_p(0, \infty)$.

The object of this paper is to apply the Erdélyi-Kober operators of fractional integration and differentiation to develope the theory of Meijer transform and to prove a number of theorems connecting Meijer transforms of different orders.

MEIJER [*5*] in 1940 defined his transform by the integral equation

$$F(s) = \int_0^\infty e^{-\frac{1}{2}st} W_{k+\frac{1}{2},m}(st)\,(st)^{-k-\frac{1}{2}} f(t)\,dt, \tag{1.1}$$

and its inverse

$$f(t) = \operatorname*{Lt}_{\lambda\to\infty} \frac{\Gamma(1-k+m)}{2\pi i\,\Gamma(1+2m)} \int_{\beta-\lambda i}^{\beta+\lambda i} e^{\frac{1}{2}st} M_{k-\frac{1}{2},m}(st)\,(st)^{k-\frac{1}{2}} F(s)\,ds \tag{1.2}$$

where $M_{k,m}(z)$ and $W_{k,m}(z)$ are the two "Whittaker Functions".

We will denote (1.1) symbolically as

$$f(t) \overset{\substack{k+\frac{1}{2}\\ m}}{\longrightarrow} F(s),$$

and $F(s)$ will be called the Meijer transform of $f(t)$.

For $k = \pm m$, (1.1) reduces to the classical Laplace transform

$$F(s) = \int_0^\infty e^{-st} f(t)\,dt.$$

R. S. VARMA [*10*] in 1951 defined a similar transform by the integral equation

$$f(s) = \int_0^\infty (st)^{m-\frac{1}{2}} e^{-\frac{1}{2}st} W_{k,m}(st)\,g(t)\,dt. \tag{1.3}$$

If we take, $m = 0$, $k = \frac{1}{2}$, the above integral reduces to Laplace transform.

K. M. SAKSENA [*8*], [*9*] later gave the inversion formulae for this transform and studied some of its properties.

If we replace k by $k + \frac{1}{2}$ and $g(t)$ by $t^{-k-m}\Phi(t)$ we find that (1.3) reduces to the well-known MEIJER's transform defined in (1.1).

For the sake of convenience I define the operators of fractional integration and differentiation,

$$I^{+}_{\eta,\alpha} f = \frac{z^{-\eta-\alpha}}{\Gamma(\alpha)} \int_0^z (z-t)^{\alpha-1} t^{\eta} f(t)\, dt, \tag{1.4}$$

$$K^{-}_{\eta,\alpha} f = \frac{z^{\eta}}{\Gamma(\alpha)} \int_z^{\infty} (t-z)^{\alpha-1} t^{-\eta-\alpha} f(t)\, dt, \tag{1.5}$$

$$I^{-}_{\eta,\alpha} f = \frac{z^{-\eta-\alpha}}{\Gamma(\alpha)} \int_z^{\infty} (t-z)^{\alpha-1} t^{\eta} f(t)\, dt, \tag{1.6}$$

$$K^{+}_{\eta,\alpha} f = \frac{z^{\eta}}{\Gamma(\alpha)} \int_0^z (z-t)^{\alpha-1} t^{-\eta-\alpha} f(t)\, dt, \tag{1.7}$$

where $\operatorname{Re}\eta > -\frac{1}{2}$, $\operatorname{Re}\alpha > 0$ and $f(t)$ belongs to the class $L_2(0, \infty)$.

2. The fundamental properties of $F(s)$ defined by (1.1) are expressed in the following two theorems:

Theorem 1. Let $1 \leq p \leq 2$ and let $f(x)$ belong to $L_p(0, \infty)$. Then the transform $F(x)$ as defined by (1.1) exists and

$$\left\{\int_0^{\infty} |F(x)|^{p'} dx\right\}^{1/p'} \leq A \left\{\int_0^{\infty} |f(x)|^{p} dx\right\}^{1/p} \quad \left(\frac{1}{p} + \frac{1}{p'} = 1\right).$$

When $p = 1$, the left hand term is to be replaced by the essential upper bound of $|F(x)|$ in $(0, \infty)$.

Proof. Since,

$$|e^{-\frac{1}{2}z} W_{k+\frac{1}{2},m}(z)\, z^{-k-\frac{1}{2}}| \leq e^{-\operatorname{Re}(z)}$$

we have,

$$|F(x)|_2 {}^{*}) \leq C \left(\int_0^N |f(t)|^2 dt\right)^{\frac{1}{2}}$$

where C is an absolute constant, and

$$|F(x)|_{\infty} = \operatorname*{ess.\,u.\,b.}_{0<x<\infty} |F(x)| \leq C \int_0^N |F(t)|\, dt.$$

Finally from the convexity theorem of M. RIESZ [7] it follows that,

$$|F(x)|_{p'} \leq C \left(\int_0^N |f(t)|^p dt\right)^{1/p}$$

which proves the theorem.

Theorem 2. Let $1 \leq p \leq 2$ and $f_j(t)$ $(j = 1, 2)$ belong to $L_p(0, \infty)$, and

$$F_j(s) = \int_0^{\infty} e^{-\frac{1}{2}st} W_{k+\frac{1}{2},m}(st)\, (st)^{-k-\frac{1}{2}} f_j(t)\, dt \qquad (j = 1, 2)$$

*) $|F(x)|_p$ means $\left|\int_0^{\infty} |F(t)|^p dt\right|^{1/p}$.

then

$$\int_0^\infty f_1 F_2 \, dx = \int_0^\infty F_1 f_2 \, dx.$$

This is PARSEVAL's Theorem. The proof is quite elementary.

3. In the following section I shall give a theorem connection Meijer Transforms of different orders.

Theorem 3. Let l be a positive integer, $\mathrm{Re}(k \pm m) < \frac{1}{p'}$ and $2\,\mathrm{Re}(k) + l - 1 < \frac{1}{p}$ and let $f(t)$ belong to $L_p(0, \infty)$ $(1 \leq p \leq 2)$ and

(a) $$f(t) \overset{k+\frac{1}{2}}{\underset{m}{\longrightarrow}} F(x).$$

Let

(3.1) $$g(z) = I_{2k-1,l} f = \frac{z^{-2k+1-l}}{\Gamma(l)} \int_z^\infty (t-z)^{l-1} f(t)\, t^{2k-1} \, dt$$

and

$$G = I_{2k-1,l} F$$

then

$$G = -\int_0^\infty e^{-st/2} (st)^{-k-\frac{1}{2}} W_{k+\frac{1}{2},m}(st)\, g(t)\, dt + \\ + \left(\frac{1-4k}{s}\right) \int_0^\infty e^{-st/2} (st)^{-k+\frac{1}{2}} W_{k-\frac{1}{2},m}(st) \left(\frac{g(t)}{t}\right) dt.$$

(b)
$$F(s) = -\int_0^\infty e^{-st/2} (st)^{-k-\frac{1}{2}} W_{k+\frac{1}{2},m}(st)\, f(t)\, dt + \\ + \left(\frac{1-4k}{s}\right) \int_0^\infty e^{-st/2} (st)^{-k+\frac{1}{2}} W_{k-\frac{1}{2},m}(st) \left(\frac{f(t)}{t}\right) dt.$$

Let

(3.2) $$g(z) = K_{2k-1,l} f = \frac{z^{2k-1}}{\Gamma(l)} \int_z^\infty (z-t)^{l-1} f(t)\, t^{-2k+1-l} \, dt$$

and

$$G = K_{2k-1,l} F; \quad \text{then} \quad g(t) \overset{k+\frac{1}{2}}{\underset{m}{\longrightarrow}} G.$$

I shall prove these theorems for $l = 1$ only. The more general statement then immediately follows by repeated application of that particular case.

Proof of Theorem 3. (a) For $l = 1$, by theorem 1 $F(x) \in L_{p'}(0, \infty)$.

Also since $2\,\mathrm{Re}\, k < \frac{1}{p}$, $g(x)$ and $G(x)$ belong to $L_p(0, \infty)$ and $L_{p'}(0, \infty)$ respectively[1]).

Now we define

$$f_1(t) = 0, \quad \text{when} \quad 0 \leq t \leq s \quad \text{and} \quad f_1(t) = t^{2k-1} \quad \text{when} \quad t > s.$$

[1]) See Theorem 330 of HARDY, LITTLEWOOD; POLYA's Inequality (Cambridge 1934).

Plainly $f_1(t) \in L_p(0, \infty)$ over every $x>0$ and

$$(3.3) \qquad \begin{cases} F_1(t) = \int_s^\infty e^{-\frac{1}{2}ut}(ut)^{-k-\frac{1}{2}} W_{k+\frac{1}{2},m}(ut)\, u^{2k-1}\, du \\ \qquad = t^{-2k} e^{-st/2}(st)^{k-\frac{1}{2}} W_{k-\frac{1}{2},m}(st). \end{cases}$$

The application of Theorem 2 to $f(t)$, $F(t)$ and $f_1(t)$, $F_1(t)$ gives

$$\int_s^\infty F(t)\, t^{2k-1}\, dt = \int_0^\infty t^{-2k} e^{-st/2}(st)^{k-\frac{1}{2}} W_{k-\frac{1}{2},m}(st)\, f(t)\, dt$$

i.e.,

$$\begin{aligned} G(s) &= s^{-2k} \int_s^\infty F(t)\, t^{2k-1}\, dt \\ &= \int_0^\infty e^{-st/2}(st)^{-k-\frac{1}{2}} W_{k-\frac{1}{2},m}(st)\, f(t)\, dt \\ &= -\int_0^\infty e^{-st/2}(st)^{-k-\frac{1}{2}} W_{k-\frac{1}{2},m}(st)\, t^{-2k+1}\, d\left[t^{2k} g(t)\right]. \end{aligned}$$

Hence

$$G(s) = \lim_{N\to\infty} G(s, N),$$

where

$$(3.4) \qquad G(s, N) = -s^{-k-\frac{1}{2}} \int_0^N e^{-st/2} t^{-3k+\frac{1}{2}} W_{k-\frac{1}{2},m}(st)\, d\left[t^{2k} g(t)\right].$$

Now, integrating by parts, we have,

$$(3.5) \qquad \begin{cases} G(s, N) = -s^{-k-\frac{1}{2}} \left[e^{-st/2} t^{-3k+\frac{1}{2}} W_{k-\frac{1}{2},m}(st)\, t^{2k} g(t)\right]_0^N + \\ \qquad + s^{-k-\frac{1}{2}} \int_0^N \frac{d}{dt}\left[e^{-st/2} t^{-3k+\frac{1}{2}} W_{k-\frac{1}{2},m}(st)\right] t^{2k} g(t)\, dt \\ \qquad = G_1(s, N) + G_2(s, N). \end{cases}$$

We have for $t\to 0$,

$$e^{-st/2} t^{-k+\frac{1}{2}} W_{k-\frac{1}{2},m}(st) \sim A\, t^{1-k-m} + B\, t^{1-k+m}$$

where A and B are constants and from the asymptotic expansion of $W_{k,m}(t)$ for $t\to\infty$

$$e^{-st/2} t^{-k+\frac{1}{2}} W_{k-\frac{1}{2},m}(st) \sim e^{-st}\left\{1 + O(t^{-1})\right\}$$

where $|\arg(t)|<\pi$, $\mathrm{Re}(s)>0$.

Also by applying HÖLDER's inequality to (2.1) for all values of t, we have,

$$|g(t)| \le t^{-2\,\mathrm{Re}(k)} \left\{\int_0^t |f(t)|^p\, dt\right\}^{1/p} \left\{\int_t^\infty u^{(2\,\mathrm{Re}(k)-1)p'}\, du\right\}^{1/p'} < A\, t^{-1/p} \qquad (p>1)$$

$$|g(t)| \le t^{-2\,\mathrm{Re}(k)} \int_t^\infty |f(t)|\, t^{2\,\mathrm{Re}(k)-1}\, dt \le t^{-1} \int_t^\infty |f(t)|\, dt \le A\, t^{-1} \qquad (p=1)$$

where A does not depend on t. Hence the expression in the square brackets in $G_1(s, N)$ vanishes for $t=0$ where $\mathrm{Re}\,(k \pm m) < \frac{1}{p'}$ and $2\,\mathrm{Re}\,k < \frac{1}{p}$. Also

$$G_1(s, N) = e^{-sN}\{N^{-1/p} + O(N^{-1-1/p})\} \to 0 \tag{3.6}$$

as $N \to \infty$ under our hypothesis on p, k, m.

Again we have,

$$\begin{cases} \frac{d}{dt}\,[e^{-st/2}\,t^{-3k+\frac{1}{2}}\,W_{k-\frac{1}{2},m}(st)] \\ \qquad = -\,s^{-k+\frac{1}{2}}\,e^{-st/2}\,(st)^{k-\frac{1}{2}}\,t^{-4k}\,[W_{k+\frac{1}{2},m}(st) - (1-4k)\,W_{k-\frac{1}{2},m}(st)] \end{cases} \tag{3.7}$$

and hence

$$\begin{cases} G_2(s, N) = -\int_0^N e^{-st/2}\,(st)^{-k-\frac{1}{2}}\,W_{k+\frac{1}{2},m}(st)\,g(t)\,dt + \\ \qquad\qquad + \left(\frac{1-4k}{s}\right)\int_0^N e^{-st/2}\,(st)^{-k+\frac{1}{2}}\,W_{k-\frac{1}{2},m}(st)\left(\frac{g(t)}{t}\right)dt \end{cases} \tag{3.8}$$

and

$$G(s) = \lim_{N\to\infty} G_2(s, N) \qquad \text{by (3.5) and (3.6).}$$

Now $g(x) \in L_p$; therefore by Theorem 1 and (3.8) the limit in the mean of index p' of $G_2(s, N)$ exists and equals

$$-\int_0^\infty e^{-st/2}\,(st)^{-k-\frac{1}{2}}\,W_{k+\frac{1}{2},m}(st)\,g(t)\,dt + \left(\frac{1-4k}{s}\right)\int_0^\infty e^{-st/2}\,(st)^{-k+\frac{1}{2}}\,W_{k-\frac{1}{2},m}(st)\left(\frac{g(t)}{t}\right)dt.$$

This completes the proof.

The proof of 3 (b) is similar to that of Theorem 3 (a) except that the integration by parts has to be done the other way round, as it were.

Instead of function f_1 we have to use here,

$$f_2(t) = t^{-2k}, \quad 0 \le t \le s; \quad f_2(t) = 0, \quad t > s.$$

Instead of (3.3) we obtain,

$$\begin{cases} F_2(t) = -\int_0^s e^{-ut/2}\,(ut)^{-k-\frac{1}{2}}\,[W_{k+\frac{1}{2},m}(ut) - (1-4k)\,W_{k-\frac{1}{2},m}(ut)]\,u^{-2k}\,du \\ \qquad = t^{-k-\frac{1}{2}}\int_0^s \frac{d}{dt}\,[e^{-ut/2}\,u^{-3k+\frac{1}{2}}\,W_{k-\frac{1}{2},m}(ut)]\,du \\ \qquad = t^{-k-\frac{1}{2}}\,e^{-st/2}\,s^{-3k+\frac{1}{2}}\,W_{k-\frac{1}{2},m}(st). \end{cases} \tag{3.9}$$

We then apply Theorem 2 to f, f_2, and F, F_2, writing in this case,

$$f(t) = t^{2k}\,\frac{d}{dt}\,[g(t)\,t^{-2k+1}] \tag{3.10}$$

and integrating by parts. In doing so, instead of (3.7) we have to use,

$$\frac{d}{dt}\left[e^{-st/2}\,t^{k-\frac{1}{2}}\,W_{k-\frac{1}{2},m}(st)\right] = -\,e^{-st/2}\,t^{k-\frac{3}{2}}\,W_{k+\frac{1}{2},m}(st).$$

The rest of the proof follows closely that of Theorem 3 (a).

4. In this section I have applied the operators of fractional integration and differentiation to deduce some very interesting properties of Meijer Transform. The results have been given in the form of theorems. The rule for fractional differentiation is

$$\frac{d^{\nu}\omega^{\mu-1}}{d\omega^{\nu}} = \frac{\Gamma(\mu)\,\omega^{\mu-\nu-1}}{\Gamma(\mu-\nu)}.$$

Theorem 4. If

$$f(s) \xrightarrow[m]{k+\frac{1}{2}} F(t)$$

then

$$\frac{\partial^{2k}}{\partial t^{2k}} f(t) \xrightarrow[m]{-k+\frac{1}{2}} s^{2k} F(s).$$

Proof. We begin with the inversion formula, (1.2) of (1.1) viz,

$$f(t) = \frac{\Gamma(1-k+m)}{2\pi i\,\Gamma(1+2m)} \int_{\beta-\infty i}^{\beta+\infty i} e^{st/2}\,M_{k-\frac{1}{2},m}(st)\,(st)^{k-\frac{1}{2}}\,F(s)\,ds$$

$$= \frac{\Gamma(1-k+m)}{2\pi i\,\Gamma(1+2m)} \int_{\beta+\infty i}^{\beta-\infty i} \left\{\sum_{r=0}^{\infty} \frac{(-k+m+1)_r}{r!\,(1+2m)_r}\,(st)^{k+m+r}\right\} F(s)\,ds.$$

Hence,

$$\frac{\partial^{\nu} f(t)}{\partial t^{\nu}} = \frac{\Gamma(1-k+m)}{2\pi i\,\Gamma(1+2m)} \frac{\partial^{\nu}}{\partial t^{\nu}} \left\{\int_{\beta-\infty i}^{\beta+\infty i} \left[\sum_{r=0}^{\infty} \frac{(-k+m+1)_r}{r!\,(1+2m)_r}\,(st)^{k+m+r}\right] F(s)\,ds\right\}.$$

As the change of integration and summation is easily justified, we can differentiate term by term under the sign of integration and hence we get after summation

$$(4.1)\quad \left\{\begin{aligned} \frac{\partial^{\nu} f(t)}{\partial t^{\nu}} = {} & \frac{\Gamma(1-k+m)}{2\pi i\,\Gamma(1+2m)} \frac{\Gamma(1+k+m)}{\Gamma(k+m+1-\nu)}\,t^{-\nu} \times \\ & \times \int_{\beta-\infty i}^{\beta+\infty i} (st)^{k+m}\,{}_2F_2\left[\begin{matrix} 1-k+m,\ 1+k+m; \\ 1+2m,\ k+m+1-\nu; \end{matrix} st\right] F(s)\,ds. \end{aligned}\right.$$

Taking $\nu = 2k$ we get,

$$\frac{\partial^{2k} f(t)}{\partial t^{2k}} = \frac{\Gamma(1+k+m)}{2\pi i\,\Gamma(1+2m)}\,t^{2k} \int_{\beta-\infty i}^{\beta+\infty i} (st)^{k+m}\,{}_1F_1[k+m+1;\,1+2m;\,st]\,F(s)\,ds$$

$$= \frac{\Gamma(1+k+m)}{2\pi i\,\Gamma(1+2m)} \int_{\beta-\infty i}^{\beta+\infty i} e^{st/2}\,(st)^{-k-\frac{1}{2}}\,M_{-k-\frac{1}{2},m}(st)\left\{s^{2k}F(s)\right\} ds.$$

Hence on inverting, we get

$$(4.2) \qquad s^{2k}F(s) = \int_0^\infty e^{-st/2}(st)^{k-\frac{1}{2}} W_{-k+\frac{1}{2},m}(st)\left\{\frac{\partial^{2k}}{\partial t^{2k}} f(t)\right\} dt$$

and hence the result.

Theorem 5. If

$$f(s) \xrightarrow[m]{k+\frac{1}{2}} F(t)$$

then

$$s^{\alpha+k-m} I^\alpha \{s^{m-k} f(s)\} \xrightarrow[m+\frac{\alpha}{2}]{k+\frac{\alpha}{2}+\frac{1}{2}} F(t),$$

where $I^\alpha f$ is the Riemann-Liouville integral of order α.

Proof. We begin with the inversion formula, i.e., (1.2)

$$f(t) = \frac{\Gamma(1-k+m)}{2\pi i\,\Gamma(1+2m)} \int_{\beta-\infty i}^{\beta+\infty i} e^{\frac{1}{2}ts} M_{k-\frac{1}{2},m}(st)\,(st)^{k-\frac{1}{2}} F(s)\,ds.$$

Hence

$$I^\alpha s^\lambda f(s) = \frac{1}{\Gamma(\alpha)} \int_0^s (s-u)^{\alpha-1} f(u)\,u^\lambda\,du$$

$$= \frac{\Gamma(1-k+m)}{\Gamma(1+2m)\,\Gamma(\alpha)\,2\pi i} \int_0^s (s-u)^{\alpha-1} u^\lambda\,du \int_{\beta-\infty i}^{\beta+\infty i} e^{\frac{1}{2}tu} M_{k-\frac{1}{2},m}(ut)\,(ut)^{k-\frac{1}{2}} F(t)\,dt.$$

Changing the order of integration, which is justified by absolute convergence, and integrating term by term, we get

$$(4.3) \qquad \begin{cases} I^\alpha s^\lambda f(s) = \dfrac{\Gamma(1-k+m)\,\Gamma(k+m+\lambda+1)\,s^{\lambda+k+m+\alpha}}{\Gamma(1+2m)\,\Gamma(k+m+\lambda+\alpha+1)\,2\pi i} \times \\ \qquad \times \displaystyle\int_{\beta-\infty i}^{\beta+\infty i} t^{k+m}\,{}_2F_2\left[\begin{matrix} 1-k+m,\ k+m+\lambda+1; \\ 1+2m,\ k+m+\lambda+\alpha+1; \end{matrix} st\right] F(t)\,dt. \end{cases}$$

Taking $\lambda = m-k$, we get,

$$I^\alpha s^{m-k} f(s) = \frac{\Gamma(1-k+m)}{2\pi i\,\Gamma(1+\alpha+2m)} \int_{\beta-\infty i}^{\beta+\infty i} t^{k+m}\,{}_1F_1[1-k+m;\,\alpha+1+2m;\,st]\,F(t)\,dt.$$

Hence

$$(4.4) \qquad \begin{cases} s^{\alpha+k-m} I^\alpha s^{m-k} f(s) = \dfrac{\Gamma(1-k+m)}{2\pi i\,\Gamma(1+\alpha+2m)} \times \\ \qquad \times \displaystyle\int_{\beta+\infty i}^{\beta-\infty i} e^{st/2}(st)^{k+\frac{\alpha}{2}-\frac{1}{2}} M_{k+\frac{\alpha}{2}-\frac{1}{2},\,m+\frac{\alpha}{2}}(st)\,F(t)\,dt. \end{cases}$$

Hence on inverting, we get,

$$F(t) = \int_0^\infty e^{-\frac{1}{2}st} W_{k+\frac{\alpha}{2}+\frac{1}{2},\, m+\frac{\alpha}{2}}(st)\,(st)^{-k-\frac{\alpha}{2}-\frac{1}{2}} \left\{ s^{\alpha+k-m} I^{\alpha}[s^{m-k} f(s)] \right\} ds,$$

which completes the proof.

Theorem 6. If

$$f(s) \xrightarrow[m]{k+\frac{1}{2}} F(t)$$

then

$$z^{\alpha} I^{+}_{m-k,\alpha} f(z) \xrightarrow[m+\frac{\alpha}{2}]{k+\frac{\alpha}{2}+\frac{1}{2}} s^{-\alpha} F(s).$$

Proof. Again starting with the inversion formula

$$f(t) = \frac{\Gamma(1-k+m)}{2\pi i\, \Gamma(1+2m)} \int_{\beta-\infty i}^{\beta+\infty i} e^{\frac{1}{2}ts} M_{k-\frac{1}{2},\, m}(st)\,(st)^{k-\frac{1}{2}} F(s)\, ds,$$

we have

$$I^{+}_{\eta,\alpha} f(z) = \frac{1}{\Gamma(\alpha)} z^{-\eta-\alpha} \int_0^z (z-t)^{\alpha-1} t^{\eta} f(t)\, dt$$

$$= \frac{1}{\Gamma(\alpha)} z^{-\eta-\alpha} \int_0^z (z-t)^{\alpha-1} t^{\eta}\, dt\, \frac{\Gamma(1-k+m)}{2\pi i\, \Gamma(1+2m)} \times$$

$$\times \int_{\beta-\infty i}^{\beta+\infty i} e^{\frac{1}{2}ts} M_{k-\frac{1}{2},\, m}(st)\,(st)^{k-\frac{1}{2}} F(s)\, ds.$$

Changing the order of integration and summation and integrating term by term, we get,

$$(4.5) \quad \begin{cases} I^{+}_{\eta,\alpha} f(z) = \dfrac{\Gamma(1-k+m)\,\Gamma(k+m+\eta+1)}{2\pi i\, \Gamma(1+2m)\,\Gamma(\alpha+k+m+\eta+1)} z^{k+m} \times \\ \qquad \times \displaystyle\int_{\beta-\infty i}^{\beta+\infty i} s^{k+m}\, {}_2F_2\left[\begin{matrix} -k+m+1,\, k+m+\eta+1; \\ 1+2m,\, k+m+\eta+\alpha+1; \end{matrix} sz \right] F(s)\, ds. \end{cases}$$

Putting $\eta = m-k$, we get,

$$(4.6) \quad \begin{cases} I^{+}_{m-k,\alpha} f(z) = \dfrac{\Gamma(1-k+m)}{2\pi i\, \Gamma(1+2m+\alpha)} \times \\ \qquad \times \displaystyle\int_{\beta-\infty i}^{\beta+\infty i} e^{\frac{1}{2}sz} (sz)^{k+\frac{\alpha}{2}-\frac{1}{2}} M_{k+\frac{\alpha}{2}-\frac{1}{2},\, m+\frac{\alpha}{2}}(sz) \left\{ (zs)^{-\alpha} F(s) \right\} ds \end{cases}$$

Hence on inverting,

$$(4.7) \quad s^{-\alpha} F(s) = \int_0^\infty e^{-\frac{1}{2}sz} W_{k+\frac{\alpha}{2}+\frac{1}{2},\, m+\frac{\alpha}{2}}(sz)\,(sz)^{-k-\frac{\alpha}{2}-\frac{1}{2}} \left\{ z^{\alpha} I^{+}_{m-k,\alpha} f(z) \right\} dz$$

which proves the theorem.

Similarly,

Theorem 7. If

$$f(s) \xrightarrow[m]{k+\frac{1}{2}} F(t)$$

then

$$s^{\alpha} K^{+}_{k-m-\alpha,\alpha} f(s) \xrightarrow[m+\frac{\alpha}{2}]{k+\frac{\alpha}{2}+\frac{1}{2}} z^{-\alpha} F(z).$$

I am grateful to Dr. R. P. AGARWAL for the ungrudging help that he has given in the preparation of this paper.

References

[1] ERDÉLYI, A.: Quart. J. Math. (Oxford) **11**, 293—303 (1940). — [2] ERDÉLYI and KOBER: Quart. J. Math. (Oxford) **11**, 212—221 (1940). — [3] KOBER, H.: Quart. J. Math. (Oxford) **11**, 193—211 (1940). — [4] KOBER, H.: Trans. Amer. Math. Soc. **50**, 160—174 (1941). — [5] MEIJER, C. S.: Proc. Amsterdam Akad. Wet. **43**, 599—608, 702—711 (1940). — [6] MEIJER, C. S.: Proc. Amsterdam Akad. Wet. **44**, 727—737, 831—839 (1940). — [7] RIESZ, M.: Acta Math. **49**, 465—497 (1927). — [8] SAKSENA, K. M.: Proc. Nat. Acad. Sci. (India) A **21**, 201—208 (1952). — [9] SAKSENA, K. M.: Proc. Nat. Inst. Sci. (India) **19**, 173—181 (1953). — [10] VARMA, R. S.: Proc. Nat. Acad. Sci. (India) A **20**, 209—216 (1951).

Department of Mathematics and Statistics, The University, Lucknow (India)

(Eingegangen am 4. August 1956)

ENDLER, O.
Math. Zeitschr. Bd. 67, S. 413—427 (1957)

Differentiation in algebraischen Funktionenkörpern von n Variablen [1])

Von
OTTO ENDLER

Die Arbeit will feststellen, wie weit die elementare Theorie der Differentiation sich auf die Untersuchung der maximaldimensionalen *(maxdim.)* Bewertungen eines algebraischen Funktionenkörpers von n Variablen *(a.F.nV.)* mit Erfolg anwenden läßt. Die von E. KÄHLER[2]) mit tieferliegenden Hilfsmitteln durchgeführten allgemeineren Untersuchungen sind bewußt unberücksichtigt geblieben.

Die Übertragung einer Reihe von Begriffen und Sätzen aus der Theorie der a.F.1V.[3]) wird dadurch ermöglicht, daß jede maxdim. Bewertung B eines a.F.nV. $\mathfrak{K}|\mathsf{K}$ für einen geeignet konstruierten Zwischenkörper l als Bewertung des a.F.1V. $\mathfrak{K}|\mathsf{l}$ aufgefaßt werden kann (§ 2). Für die Konstruktion von l sind insbesondere solche algebraisch-unabhängige *(alg.-unabh.)* n-Tupel $\mathfrak{x}=\langle x_1, \ldots, x_n\rangle$ aus $\mathfrak{B}$ von Bedeutung, aus deren Restklassen ein in $\Lambda|\mathsf{K}$ alg.-unabh. bzw. separierendes $(n-1)$-Tupel gewählt werden kann ($\mathfrak{B}$ bzw. Λ bezeichnet den zu B gehörigen Bewertungsring bzw. Restklassenkörper); n-Tupel dieser Art werden B-zulässig bzw. B-regulär genannt. B-zulässige $\mathfrak{x}$ sind dadurch ausgezeichnet, daß B über einer Polynombewertung von $\mathsf{K}[x_1, \ldots, x_n]$ liegt.

Der Zusammenhang zwischen den maxdim. Bewertungen und den Differentiationen von $\mathfrak{K}|\mathsf{K}$ wird hergestellt durch den (in Verallgemeinerung der „uniformisierenden Variablen" gebildeten) Begriff des uniformisierenden n-Tupels (§ 3): ein in $\mathfrak{K}|\mathsf{K}$ separierendes $\mathfrak{x}$ aus $\mathfrak{B}$ heißt B-uniformisierend, wenn $D^i_{\mathfrak{x}}\mathfrak{B}\subseteq\mathfrak{B}$ $(1\leq i\leq n)$ ist ($D^1_{\mathfrak{x}}, \ldots, D^n_{\mathfrak{x}}$ bezeichnen die partiellen Differentiationen nach $\mathfrak{x}$). $\mathfrak{y}$ aus $\mathfrak{B}$ ist genau dann ebenfalls B-uniformisierend, wenn die Funktionaldeterminante $|D^i_{\mathfrak{x}} y_j|_{1\ldots n}$ eine Einheit in $\mathfrak{B}$ ist. Es wird die Existenz B-uniformisierender n-Tupel bewiesen, sowie der für das Folgende grundlegende Satz: „Ist $\Lambda|\mathsf{K}$ separabelerzeugbar, so ist $\mathfrak{x}$ genau dann B-uniformisierend, wenn $\mathfrak{x}$ B-regulär und B unverzweigt bezüglich $\mathsf{K}(x_1, \ldots, x_n)$ ist."

Als erste Anwendung dieses Satzes wird in § 4 der Begriff „uniformisierend" auf gestufte Bewertungen $\boldsymbol{B}=\langle B^{\langle 1\rangle}, \ldots, B^{\langle n\rangle}\rangle$ eines a.F.nV. $\mathfrak{K}^{\langle n\rangle}|\mathsf{K}$ übertragen ($B^{\langle i\rangle}$ maxdim. Bewertung von $\mathfrak{K}^{\langle i\rangle}|K$, $\mathfrak{K}^{\langle i-1\rangle}=\Lambda^{\langle i\rangle}$), und es wird

[1]) Umgearbeitete Fassung meiner Dissertation (Bonn 1955).

[2]) KÄHLER, E.: Algebra und Differentialrechnung, Ber. Math. Tagung Berlin 1953, S. 58—163.

[3]) Vgl. C. CHEVALLEY: Introduction to the Theory of Algebraic Functions of one Variable. Math. Surveys VI. New York 1951. Zitiert mit „CH.".

insbesondere gezeigt, wie er mit den Differentiationen in den durch $\boldsymbol{B}$ definierten allgemeinen Bewertungsringen $\mathfrak{C}_1 \subset \mathfrak{C}_2 \subset \cdots \subset \mathfrak{C}_n = \mathfrak{B}^{\langle n \rangle}$ zusammenhängt.

In § 5 (der nur §§ 1–3 voraussetzt) werden auf die (alternierenden) Differentiale n-ten Grades eines separabelerzeugbaren a.F.nV. $\mathfrak{R}|\mathsf{K}$ dessen maxdim. Bewertungen so übertragen, daß im Falle $n=1$ Übereinstimmung mit CH. besteht; hierzu braucht man, wie ein Beispiel zeigt, eine Verallgemeinerung der Differente auf transzendente Erweiterungen. Anschließend wird der für $n=1$ bekannte Zusammenhang zwischen Differential und Differente (Riemannsche Formel): $B(dx) = \mathrm{Diff}_{\mathfrak{R}\,\mathsf{K}(x)} B$ ($x \in \mathfrak{B}$, separierend) auf separierende B-zulässige n-Tupel $\mathfrak{x}$ verallgemeinert, allerdings nur als eine Ungleichung, in der das Gleichheitszeichen genau dann gilt, wenn $\mathfrak{x}$ „B-semiregulär" ist (dies trifft stets zu für $n=1$, und für beliebiges n im Falle der Charakteristik 0). Ein Beispiel eines separierenden B-zulässigen, aber nicht B-semiregulären n-Tupels zeigt, daß die hier entwickelten Hilfsmittel nicht ausreichen, eine befriedigende Verallgemeinerung der Riemannschen Formel auf a.F.nV. zu gewinnen[4]).

Schließlich werden in § 6 die Ergebnisse von §§ 2–3 zur Untersuchung der projektiven Bewertungssysteme[5]) verwendet. Insbesondere wird gezeigt, daß bei vollkommenem Koeffizientenkörper K zwei beliebige (zu separierenden n-Tupeln gehörige) projektive Bewertungssysteme sich höchstens um endlich viele Bewertungen unterscheiden.

§ 1. Bezeichnungen

Die Buchstaben K, $\mathfrak{k}$, $\mathfrak{R}$, λ, Λ, $\mathfrak{l}$, $\mathfrak{L}$ bezeichnen stets (kommutative) Körper; $Ch(\mathsf{K}) =$ Charakteristik von K, $\mathsf{K}^{\cdot} = \mathsf{K} - \{0\}$, $\mathsf{K}_{\mathfrak{R}} = \{a \in \mathfrak{R} \mid a$ algebraisch über $\mathsf{K}\}$ ($\mathsf{K} \subseteq \mathfrak{R}$), $\mathsf{K}(\mathfrak{x}) = \mathsf{K}(x_1, \ldots, x_n)$. Eine endlicherzeugbare Körpererweiterung $\mathfrak{R}|\mathsf{K}$ vom Transzendenzgrad *(Trgr.)* $n>0$ heißt ein *a.F.nV.* mit dem Koeffizientenkörper K und Konstantenkörper $\mathsf{K}_{\mathfrak{R}}$; ist K vollkommen, so ist $\mathfrak{R}|\mathsf{K}$ separabelerzeugbar.

Die Ableitung eines Polynoms $F = F(X_1, \ldots, X_n)$ nach X_j wird mit $F^{(j)}$ bezeichnet; es ist $F^{(j)} \neq 0$, genau wenn F separabel (j) (d.h. separabel in X_j) ist.

Ein n-Tupel $\mathfrak{x} = \langle x_1, \ldots, x_n \rangle$ aus $\mathfrak{R}$ heißt *minimalabhängig* (bezüglich K), wenn $\mathfrak{x}$ alg.-abh., aber für ein j $\mathfrak{x}^{(j)} = \langle x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n \rangle$ alg.-unabh. bezüglich K ist; insbesondere heißt dann $\mathfrak{x}$ *algebraisch* (j) und, falls x_j sogar separabel über $\mathsf{K}(\mathfrak{x}^{(j)})$ ist, *separabel* (j). Zu $\mathfrak{x}$ gehört ein (bis auf einen Faktor $\in \mathsf{K}^{\cdot}$ eindeutig bestimmtes) irreduzibles Polynom $P = P(X_1, \ldots, X_n)$ über K mit $P(\mathfrak{x}) = 0$. Jedes Polynom F über K mit $F(\mathfrak{x}) = 0$ ist durch P teilbar; $\mathfrak{x}$ ist separabel (j) bzw. algebraisch (j), genau wenn P in X_j separabel ist bzw. einen Grad $f_j > 0$ besitzt (dann ist $f_j = [\mathsf{K}(\mathfrak{x}) : \mathsf{K}(\mathfrak{x}^{(j)})]$). Ist K vollkommen, so ist jedes minimalabhängige $\mathfrak{x}$ aus $\mathfrak{R}$ separabel (j) für ein j.

---

[4]) Inzwischen hat F. K. SCHMIDT die Riemannsche Formel für a.F.nV. in voller Allgemeinheit bewiesen (noch unveröffentlicht).

[5]) Vgl. W. KRULL: Über geschlossene Bewertungssysteme. J. f. Math. **190**, 75–92 (1952). Zitiert mit KR.

Abbildungen D von $\mathfrak{R}$ in $\mathfrak{L} \supseteq \mathfrak{R}$, für die $D(a_1+a_2) = Da_1 + Da_2$ und $D(a_1 \cdot a_2) = a_1 \cdot Da_2 + a_2 \cdot Da_1$ $(a_1, a_2 \in \mathfrak{R})$ gilt, heißen *Differentiationen von* $\mathfrak{R}$ *in* $\mathfrak{L}$, insbesondere *von* $\mathfrak{R}|\mathsf{K}$ *in* $\mathfrak{L}$, wenn $D\mathsf{K} = \{0\}$ ist. Ihre Gesamtheit bildet ersichtlich einen $\mathfrak{L}$-Modul $\boldsymbol{D}_{\mathfrak{R},\mathfrak{L}}$ bzw. $\boldsymbol{D}_{\mathfrak{R}\mathsf{K},\mathfrak{L}}$: $(c_1 \cdot D_1 + c_2 \cdot D_2)a = c_1 \cdot D_1 a + c_2 \cdot D_2 a$ $(a \in \mathfrak{R};\ c_1, c_2 \in \mathfrak{L})$. Ist $\mathfrak{R}|\mathsf{K}$ endlicherzeugbar vom Trgr. $n \geq 0$, $\mathfrak{x}$ separierend in $\mathfrak{R}|\mathsf{K}$ und $\mathfrak{L} \supseteq \mathfrak{R}$, so läßt sich jede Differentiation $\Delta \in \boldsymbol{D}_{\mathsf{K},\mathfrak{L}}$ auf genau eine Weise zu einer Differentiation $D \in \boldsymbol{D}_{\mathfrak{R},\mathfrak{L}}$ so fortsetzen, daß $Dx_i = y_i$ $(1 \leq i \leq n)$ für vorgegebenes $\mathfrak{y}$ aus $\mathfrak{L}$ ist[6]). Die durch $D^j_{\mathfrak{x}} x_i = \delta_{ij}$ $(1 \leq i \leq n)$ definierte Differentiation $D^j_{\mathfrak{x}} \in \boldsymbol{D}_{\mathfrak{R}\mathsf{K},\mathfrak{R}}$ heißt die *Differentiation* (j) *nach* $\mathfrak{x}$ *von* $\mathfrak{R}|\mathsf{K}$; sie ist sogar $\in \boldsymbol{D}_{\mathfrak{R}\mathsf{K}(\mathfrak{x}^{(0)}),\mathfrak{R}}$. Jedes $D \in \boldsymbol{D}_{\mathfrak{R}\mathsf{K},\mathfrak{L}}$ läßt sich als Summe $D = \sum_1^n Dx_j \cdot D^j_{\mathfrak{x}}$ darstellen, $\langle D^1_{\mathfrak{x}}, \ldots, D^n_{\mathfrak{x}} \rangle$ ist also eine Basis von $\boldsymbol{D}_{\mathfrak{R}\mathsf{K},\mathfrak{L}}$. Man zeigt leicht: Ein n-Tupel $\mathfrak{y}$ aus $\mathfrak{R}$ ist genau dann ebenfalls separierend in $\mathfrak{R}|\mathsf{K}$, wenn die Funktionaldeterminante $D_{\mathfrak{x}}\mathfrak{y} = |D^j_{\mathfrak{x}} y_i|_{1\ldots n} \neq 0$ ist.

Wir betrachten nur diskrete normierte *Bewertungen* B von $\mathfrak{R}$[7]), d.h. Abbildungen von $\mathfrak{R}^\cdot$ *auf* die Menge aller ganzen Zahlen mit $B(a_1 \cdot a_2) = B(a_1) + B(a_2)$ und $B(a_1 + a_2) \geq \min\{B(a_1), B(a_2)\}$ $(a_1, a_2 \in \mathfrak{R})$, und wir setzen $B(0) = \infty$. $\mathfrak{B}$, $\mathfrak{M}$, Λ, $\mathfrak{R}^*$, B^*, Φ bezeichnen den zugehörigen Bewertungsring bzw. das Bewertungsideal von $\mathfrak{B}$ bzw. den Restklassenkörper $\mathfrak{B}/\mathfrak{M}$ bzw. die B-perfekte Hülle von $\mathfrak{R}$ bzw. die Fortsetzung von B auf $\mathfrak{R}^*$ bzw. die Restklassenbildung mod $\mathfrak{M}$ (und mod $\mathfrak{M}^*$; Λ und Λ^* werden wie üblich identifiziert). Es sei $\mathfrak{k} \subseteq \mathfrak{R}$; ist $\mathfrak{k} \subseteq \mathfrak{B}$, so ist B eine Bewertung von $\mathfrak{R}|\mathfrak{k}$ (d.h. $B(\mathfrak{k}^\cdot) = \{0\}$), anderenfalls liegt B über der Bewertung b von $\mathfrak{k}$ mit dem Bewertungsring $\mathfrak{b} = \mathfrak{B} \cap \mathfrak{k}$. Im ersteren Falle wird $\mathfrak{k}$ durch Φ isomorph in Λ abgebildet (falls $\mathfrak{k} = \mathsf{K}$ als Koeffizientenkörper angesehen wird, identifiziert man $\Phi\mathsf{K}$ mit K), im letzteren Falle wird λ (wegen $\mathfrak{m} = \mathfrak{M} \cap \mathfrak{k}$) in Λ eingebettet $(\varphi a = \Phi a\ (a \in \mathfrak{b}))$. Der letztgenannte Fall trifft stets zu, wenn $\mathfrak{R}|\mathfrak{k}$ endlich-algebraisch ist; dann liegen über jeder Bewertung b von $\mathfrak{k}$ nur endlich viele Bewertungen $B = B^{(1)}, \ldots, B^{(r)}$ $(r > 0)$. $e = e_{\mathfrak{R}\mathfrak{k}} = B(a) \cdot (b(a))^{-1}$ $(a \in \mathfrak{k}^\cdot)$ bzw. $g = g_{\mathfrak{R}\mathfrak{k}} = [\Lambda : \lambda]$ $(< \infty)$ bezeichnet den *Verzweigungsindex* bzw. *Relativgrad von* B *bezüglich* $\mathfrak{k}$; ist $e = 1$, so heißt B *unverzweigt bezüglich* $\mathfrak{k}$, ist $\Lambda|\lambda$ separabel, so heißt B *separabel bezüglich* $\mathfrak{k}$.

Eine Bewertung B des a.F.nV. $\mathfrak{R}|\mathsf{K}$ heißt *maxdim.* bezüglich K, wenn $\Lambda|\mathsf{K}$ den höchstmöglichen Trgr. $n-1$ besitzt.

Ein Bewertungssystem $\mathfrak{S}$ von $\mathfrak{R}$ heißt *streng unabhängig*, wenn zu jedem endlichen Tupel $\langle B_1, \ldots, B_s \rangle$ paarweise verschiedener Bewertungen $\in \mathfrak{S}$, jedem s-Tupel $\langle n_1, \ldots, n_s \rangle$ ganzer Zahlen und jedem $\langle y_1, \ldots, y_s \rangle$ aus $\mathfrak{R}$ ein $y \in \mathfrak{R}$ existiert, so daß $B_i(y - y_i) = n_i$ $(1 \leq i \leq s)$ und $B(y) \geq 0$ $(B \in \mathfrak{S} - \{B_1, \ldots, B_s\})$ ist. Gilt dies wenigstens für $y_1 = \cdots = y_s = 0$, so heißt $\mathfrak{S}$ *unabhängig*.

Jedes endliche Bewertungssystem ist bekanntlich streng unabhängig.

---

[6]) Vgl. N. Bourbaki: Algèbre, Paris 1950, V § 9 N° 1 Prop. 4—5; für $n = 0$ ist „$\mathfrak{x}$ separierend" zu ersetzen durch „$\mathfrak{R}|\mathsf{K}$ separabel".

[7]) Nur in § 4 werden außerdem „gestufte Bewertungen" behandelt.

§ 2. B-zulässige und B-reguläre n-Tupel

Ist B eine Bewertung des a.F.nV. $\mathfrak{K}|\mathsf{K}$, $\mathfrak{x}$ aus $\mathfrak{V}$ eine Transzendenzbasis *(Tr.-B.)* von $\mathfrak{K}|\mathsf{K}$, so ist $\mathfrak{p}=\mathfrak{M}\cap\mathsf{K}[\mathfrak{x}]$ ein Primideal in $\mathsf{K}[\mathfrak{x}]$ $((0)<\mathfrak{p}<\mathsf{K}[\mathfrak{x}])$, und der Quotientenring $\mathsf{K}[\mathfrak{x}]_{\mathfrak{p}}$ ist in dem Bewertungsring $\mathfrak{b}$ der unter B liegenden Bewertung b von $\mathsf{K}(\mathfrak{x})$ enthalten. Insbesondere ist $\mathsf{K}[\mathfrak{x}]_{\mathfrak{p}}=\mathfrak{b}$, genau wenn $\mathfrak{p}$ ein minimales Primideal (also Primhauptideal) von $\mathsf{K}[\mathfrak{x}]$ ist, und dies trifft genau dann zu, wenn $\Phi\mathfrak{x}=\langle\Phi x_1,\ldots,\Phi x_n\rangle$ minimalabhängig ist. Wir nennen $\mathfrak{x}$ in diesem Falle *B-zulässig*; insbesondere *B-zulässig* (j) bzw. *B-regulär* (j), wenn $\Phi\mathfrak{x}$ algebraisch (j) bzw. $\Phi\mathfrak{x}^{(j)}$ separierend in $\Lambda|\mathsf{K}$ ist.

Eine notwendige Bedingung für die Existenz B-zulässiger bzw. B-regulärer n-Tupel ist ersichtlich, daß B maxdim. ist bzw. außerdem $\Lambda|\mathsf{K}$ separabelerzeugbar ist. Daß diese Bedingung auch hinreicht, folgt aus:

(2.1) a) *Zu je endlich vielen maxdim. Bewertungen $B_1,\ldots,B_s$ des a.F.nV. $\mathfrak{K}|\mathsf{K}$ existiert ein n-Tupel $\mathfrak{z}$ aus $\mathfrak{V}_1\cap\cdots\cap\mathfrak{V}_s$ derart, daß $B_i(z_1)=1$ und $\Phi_i\mathfrak{z}^{(1)}$ eine [falls $\Lambda_i|\mathsf{K}$ separabelerzeugbar ist, separierende] Tr.-B. von $\Lambda_i|\mathsf{K}$ ist $(1\leq i\leq s)$. Jedes solche $\mathfrak{z}$ ist B_i-zulässig (1) [bzw. sogar B_i-regulär (1)], und B_i ist unverzweigt bezüglich $\mathsf{K}(\mathfrak{z})$ $(1\leq i\leq s)$.*

b) *Ist $\mathfrak{K}|\mathsf{K}$ separabelerzeugbar, so existiert ein separierendes $\mathfrak{z}$ dieser Art.*

Beweis. a) Es sei $\langle\zeta_{i2},\ldots,\zeta_{in}\rangle$ eine (gegebenenfalls separierende) Tr.-B. von $\Lambda_i|\mathsf{K}$ $(1\leq i\leq s)$. Wegen der strengen Unabhängigkeit von $\{B_1,\ldots,B_s\}$ gibt es ein $\mathfrak{z}$ aus $\mathfrak{V}_1\cap\cdots\cap\mathfrak{V}_s$ mit $B_i(z_1)=1$ und $\Phi_i z_j=\zeta_{ij}$ $(1\leq i\leq s,\ 2\leq j\leq n)$. $\mathfrak{z}^{(1)}$ ist alg.-unabh. und B_i eine Bewertung von $\mathfrak{K}|\mathsf{K}(\mathfrak{z}^{(1)})$; wegen $B_i(z_1)>0$ ist z_1 transzendent über $\mathsf{K}(\mathfrak{z}^{(1)})$, also $\mathfrak{z}$ eine Tr.-B. von $\mathfrak{K}|\mathsf{K}$. Der Rest ist trivial.

b) Wir zeigen zunächst: Es gibt ein separierendes $\mathfrak{y}$ mit $B_i(\mathfrak{y})=\min\{B_i(y_1),\ldots,B_i(y_n)\}>1$ $(1\leq i\leq s)$. Beweis durch Induktion: Es sei $\mathfrak{u}$ separierend, $B_i(\mathfrak{u})>1$ $(1\leq i<m)$. Dann ist $\mathfrak{v}$, definiert durch $v_i=u_i$ bzw. $u_i\cdot(u_i-1)^{-1}$ für $B_m(u_i)=$ bzw. $\neq 0$, ebenfalls separierend, und $B_i(\mathfrak{v})>1$ $(1\leq i<m)$, $B_m(\mathfrak{v})\geq 0$. Für ein irreduzibles P ist $B_m\big(P(\mathfrak{v})\big)>0$; $\mathfrak{w}=\langle P(\mathfrak{v})\cdot v_1, v_2,\ldots,v_n\rangle$ ist wegen $D_{\mathfrak{v}}\mathfrak{w}=P(\mathfrak{v})+v_1\cdot P^{(1)}(\mathfrak{v})\neq 0$ separierend, ebenso $\mathfrak{x}=w_1^k\cdot\mathfrak{w}$ (für geeignetes $k>1$), und es ist $B_i(\mathfrak{x})>1$ $(1\leq i\leq m)$.

Es sei $\mathfrak{z}$ gemäß a) konstruiert; dann ist die Determinante $|D_{\mathfrak{y}}^i z_j+X_{ij}|_{1\ldots n}$ ein Polynom in den Unbestimmten X_{ij}, in jeder vom Grade 1, also ist für eine Matrix $\langle\alpha_{ij}\rangle_{1\ldots n}$ aus K $|D_{\mathfrak{y}}^i z_j+\alpha_{ij}|_{1\ldots n}\neq 0$, also ist $\Big\langle z_1+\sum_1^n{}_i\,\alpha_{i1}\cdot y_i,\ldots, z_n+\sum_1^n{}_i\,\alpha_{in}\cdot y_i\Big\rangle$ separierend und hat die gewünschte Eigenschaft.

Im folgenden sei B bzw. b stets eine maxdim. Bewertung des a.F.nV. $\mathfrak{K}|\mathsf{K}$ bzw. $\mathfrak{k}|\mathsf{K}$. Es sei $\mathfrak{x}$ eine Tr.-B. von $\mathfrak{K}|\mathsf{K}$; die Bewertungen B, für die $\mathfrak{x}$ B-zulässig bzw. B-regulär ist, lassen sich leicht übersehen:

(2.2) *Eine Tr.-B. $\mathfrak{x}$ aus $\mathfrak{V}$ ist B-zulässig und P das zu $\Phi\mathfrak{x}$ gehörige Polynom, genau wenn B über der durch $b_{\mathfrak{x},P}\big(P(\mathfrak{x})\big)=1$, $b_{\mathfrak{x},P}\big(Q(\mathfrak{x})\big)=0$ (Q teilerfremd zu P) definierten Polynombewertung $b_{\mathfrak{x},P}$ von $\mathsf{K}(\mathfrak{x})|\mathsf{K}$ liegt. In diesem Falle ist $\lambda_{\mathfrak{x},P}=\mathsf{K}(\Phi\mathfrak{x})$, $\Lambda|\mathsf{K}$ endlicherzeugbar.*

Insbesondere ist $\mathfrak{x}$ *B-regulär* (j), *genau wenn* $\Phi\mathfrak{x}$ *separabel* (j) *und* B *separabel bezüglich* $\mathsf{K}(\mathfrak{x})$ *ist.*

Hieraus folgt unmittelbar:

(2.3) *Ist* $\mathfrak{z}$ *B-zulässig* (j) [*bzw. B-regulär* (j)], *so ist für* $\mathfrak{l}=\mathsf{K}(\mathfrak{z}^{(j)})$ B *eine Bewertung des a.F.* $1V$. $\mathfrak{R}|\mathfrak{l}$ [*bzw. außerdem* $\Lambda|\Phi\mathfrak{l}$ *separabel*].

(2.3) in Verbindung mit (2.1) ermöglicht die Übertragung einer Reihe von Sätzen aus der Theorie der a.F.1V. auf a.F.nV. Wir erwähnen hier nur:

(2.4) *Ist* $\mathfrak{R}|\mathfrak{k}$ *endlich-algebraisch,* $[\mathfrak{R}:\mathfrak{k}]=k$, *so gilt*:

a) $\mathfrak{B}^{(1)}\cap\cdots\cap\mathfrak{B}^{(r)}$ *ist gleich dem Ring* $\mathfrak{J}$ *aller von* $\mathfrak{b}$ *ganzabhängigen Elemente* $\in\mathfrak{R}$.

b) $e^{(1)}\cdot g^{(1)}+\cdots+e^{(r)}\cdot g^{(r)}=k$.

c) *Zu jeder Basis* $\langle v_1,\ldots,v_k\rangle$ *von* $\mathfrak{R}|\mathfrak{k}$ *existiert ein* $s\geqq 0$ *derart, daß aus* $\sum_1^k a_i\cdot v_i\in\mathfrak{J}$ $(a_i\in\mathfrak{k})$ *folgt:* $b(a_i)\geqq -s$ $(1\leqq i\leqq k)$.

d) *Es gibt eine b-Ganzheitsbasis* $\langle w_1,\ldots,w_k\rangle$ *von* $\mathfrak{R}|\mathfrak{k}$ (*d.h.* $\mathfrak{J}=\mathfrak{b}\cdot w_1+\cdots+\mathfrak{b}\cdot w_k$).

e) *Ist* $\mathfrak{R}|\mathfrak{k}$ *normal, so ist* $\mathfrak{B}^{(i)}=\Theta_i\mathfrak{B}^{(1)}$ *für einen Automorphismus* Θ_i *von* $\mathfrak{R}|\mathfrak{k}$ $(1\leqq i\leqq r)$; *ferner ist* $e^{(i)}=e$, $g^{(i)}=g$ $(1\leqq i\leqq r)$ *also* $r\cdot g\cdot e=k$.

Beweis. Vgl. Ch. IV § 1 Theorem 1, § 2 Th. 2, § 3 Lemma 2, Th. 3 und Lemma 4.

(2.5) *Ist* $\mathfrak{z}$ *B-regulär* (j), $\mathfrak{l}=\mathsf{K}(\mathfrak{z}^{(j)})$, *so gilt*:

a) *Es existiert genau ein Repräsentantenkörper* $\mathfrak{L}^*$ *von* Λ *mit* $\mathfrak{l}_{\mathfrak{R}}\subseteqq\mathfrak{L}^*\subseteqq\mathfrak{R}^*$.

Ist $B(u)=1$, *so läßt sich jedes* $a\in\mathfrak{R}^*$ *eindeutig als Potenzreihe* $a=\sum_{-\infty}^{\infty}{}_i\, c_i\cdot u^i$ [$c_i\in\mathfrak{L}$, $c_j=0$ *bzw.* $\neq 0$ *für* $j<$ *bzw.* $=B(a)$] *darstellen.*

b) *Jede Differentiation* D *von* $\mathfrak{R}$ *mit* $D\mathfrak{l}_{\mathfrak{R}}\subseteqq\mathfrak{l}_{\mathfrak{R}}$ *läßt sich auf genau eine Weise zu einer stetigen*[8]) *Differentiation* D^* *von* $\mathfrak{R}^*$ *fortsetzen; es ist* $D^*\mathfrak{L}^*\subseteqq\mathfrak{L}^*$ *und*

$$D^*a=\sum_{-\infty}^{\infty}{}_i(D^*c_i)\cdot u^i+Du\cdot\sum_{-\infty}^{\infty}{}_i\, i\cdot c_i\cdot u^{i-1}.$$

Beweis. Vgl. Ch. III § 3 und VI § 4 (5).

(2.6) a) *Ist* $\mathfrak{R}|\mathfrak{k}$ *endlich-algebraisch,* B *separabel bezüglich* $\mathfrak{k}$ *und* e *durch* $Ch(\mathfrak{R})$ *unteilbar, so ist* $\mathfrak{R}|\mathfrak{k}$ *separabel.*

b) *Ist* $\Lambda|\mathsf{K}$ *separabelerzeugbar, so ist jedes B-reguläre* $\mathfrak{x}$ *derart, daß* $e_{\mathfrak{R}\,\mathsf{K}(\mathfrak{x})}$ *durch* $Ch(\mathfrak{R})$ *unteilbar ist, separierend in* $\mathfrak{R}|\mathsf{K}$, *also* $\mathfrak{R}|\mathsf{K}$ *separabelerzeugbar.*

Beweis. Für $Ch(\mathfrak{R})=0$ trivial. Es sei $Ch(\mathfrak{R})=p>0$. a) Ist $\mathfrak{R}|\mathfrak{k}$ inseparabel, so existiert ein $\mathfrak{k}'\supseteqq\mathfrak{k}$ derart, daß $\mathfrak{R}|\mathfrak{k}'$ Wurzelkörper vom Grade p ist. Nach (2.4) e) ist B die einzige über b' liegende Bewertung und $p=e'\cdot g'$. Ist e durch p unteilbar, so auch e', also ist $e'=1$, $g'=p$. Also ist für ein $v\in\mathfrak{B}$ $\Lambda=\lambda'(\Phi v)>\lambda'$; wegen $v^p\in\mathfrak{k}'$ ist $\Lambda|\lambda'$ inseparabel, also auch $\Lambda|\lambda$. — b) Nach (2.1) existiert ein $\mathfrak{x}$ dieser Art; aus (2.2) und a) folgt, daß $\mathfrak{R}|\mathsf{K}(\mathfrak{x})$ separabel ist.

[8]) Das heißt: für jede B^*-konvergente Folge $\langle z_i\rangle_{m\leqq i<\infty}$ aus $\mathfrak{R}^*$ ist $D^*(\lim_{i\to\infty} z_i)=\lim_{i\to\infty} D^*z_i$.

§ 3. B-uniformisierende n-Tupel

Im a.F.1V. $\mathfrak{K}|\mathsf{K}$ werden als „B-uniformisierende Variable" bekanntlich die Primelemente z von $\mathfrak{B}$ ($B(z)=1$) bezeichnet; sie sind, wenn $\Lambda|\mathsf{K}$ separabel ist, separierend in $\mathfrak{K}|\mathsf{K}$, und es gilt $D_z\mathfrak{B}\subseteqq\mathfrak{B}$. In Verallgemeinerung dieser Eigenschaft definieren wir: Ist B eine Bewertung des a.F.nV. $\mathfrak{K}|\mathsf{K}$, so heißt ein n-Tupel $\mathfrak{z}$ aus $\mathfrak{B}$ *B-uniformisierend*, wenn $\mathfrak{z}$ separierend und

$$D_{\mathfrak{z}}\mathfrak{B}\left(=\bigcup_{i=1\ldots n}D^i_{\mathfrak{z}}\mathfrak{B}\right)\subseteqq\mathfrak{B}$$

ist. (Wir verlangen nicht, daß ein $z_i\in\mathfrak{M}$ ist!)

(3.1) *Für jedes $\mathfrak{v}$ aus $\mathfrak{B}$ ist $D_{\mathfrak{z}}\mathfrak{v}\in\mathfrak{B}$; insbesondere ist $B(D_{\mathfrak{z}}\mathfrak{v})=0$, genau wenn $\mathfrak{v}$ ebenfalls B-uniformisierend ist.*

Beweis. $D_{\mathfrak{z}}\mathfrak{v}\in\mathfrak{B}$ gilt trivial; ebenso $D_{\mathfrak{v}}\mathfrak{z}=(D_{\mathfrak{z}}\mathfrak{v})^{-1}\in\mathfrak{B}$, falls auch $\mathfrak{v}$ B-uniformisierend ist. — Ist $B(D_{\mathfrak{z}}\mathfrak{v})=0$, so ist $\mathfrak{v}$ separierend, und aus $\sum_1^n D^i_{\mathfrak{z}}v_j\cdot D^j_{\mathfrak{v}}z_k=\delta_{ik}$ folgt $D^j_{\mathfrak{v}}z_k\cdot D_{\mathfrak{z}}\mathfrak{v}\in\mathsf{K}[\ldots D^i_{\mathfrak{z}}v_l\ldots]\subseteqq\mathfrak{B}$, also auch $D^j_{\mathfrak{v}}z_k\in\mathfrak{B}$ $(j,k=1,\ldots,n)$; daher ist für jedes $a\in\mathfrak{B}$ $D^j_{\mathfrak{v}}a=\sum_1^n D^j_{\mathfrak{v}}z_k\cdot D^k_{\mathfrak{z}}a\in\mathfrak{B}$.

Im folgenden sei B wieder eine maxdim. Bewertung eines a.F.nV. $\mathfrak{K}|\mathsf{K}$. Für den Existenzbeweis B-uniformisierender n-Tupel brauchen wir die folgende Verallgemeinerung eines Chevalleyschen Lemmas (CH. VI § 4, Lemma 4):

(3.2) *Zu jeder Differentiation D von $\mathfrak{K}$ mit $D\mathsf{K}\subseteqq\mathsf{K}$ gibt es ein $m\geqq 0$, so daß für alle $a\in\mathfrak{K}$ $B(Da)\geqq B(a)-m$ ist.*

Beweis. Wir wählen ein B-zulässiges $\mathfrak{x}$ mit $B(x_1)=1$ [(2.1)!], eine b-Ganzheitsbasis $\langle v_1,\ldots,v_k\rangle$ von $\mathfrak{K}|\mathsf{K}(\mathfrak{x})$ [(2.4) d)!] und ein $u\in\mathfrak{K}$ mit $B(u)=0$ und $B^{(i)}(u)>0$ $[2\leqq i\leqq r$; $b=b_{\mathfrak{x},P}$ die unter B liegende Bewertung von $\mathsf{K}(\mathfrak{x})$, $B=B^{(1)}]$. Ist $a\in\mathfrak{K}^{\cdot}$, $B(a)=t$, so ist für genügend großes s $d=a\cdot x_1^{-t}\cdot u^s\in\mathfrak{J}$, also $=\sum_1^k c_i\cdot v_i$ für gewisse $c_i\in\mathfrak{b}$ $(1\leqq i\leqq k)$, und $B(d)=0$. Hieraus folgt $a^{-1}\cdot Da=x_1^{-1}\cdot t\cdot Dx_1-u^{-1}\cdot s\cdot Du+d^{-1}\cdot\left(\sum_1^k c_i\cdot Dv_i+v_i\cdot Dc_i\right)$; da ferner wegen $\mathfrak{b}=\mathsf{K}[\mathfrak{x}]_{(P(\mathfrak{x}))}$ und $D\mathsf{K}\subseteqq\mathsf{K}$ für jedes $c\in\mathfrak{b}$ $B(Dc)\geqq\min\{0,B(Dx_1),\ldots,B(Dx_n)\}$ gilt, kann man $-m=\min\{0,B(Du),B(Dv_i),B(Dx_j)-1\,|\,1\leqq i\leqq k,\ 1\leqq j\leqq n\}$ setzen.

Hieraus folgt der Existenzsatz:

(3.3) Satz. *Ist $\mathfrak{K}|\mathsf{K}$ separabelerzeugbar, so existiert ein B-uniformisierendes $\mathfrak{z}$.*

Beweis. Es sei $\mathfrak{x}$ separierend in $\mathfrak{B}$. Nach (3.2) existiert ein m, so daß $B(D_{\mathfrak{x}}\mathfrak{v})\geqq -m$ für alle $\mathfrak{v}$ aus $\mathfrak{B}$ ist, also gibt es ein $\mathfrak{z}$ aus $\mathfrak{B}$, so daß $B(D_{\mathfrak{x}}\mathfrak{z})$ minimal ist; $\mathfrak{z}$ ist ersichtlich separierend. Für jedes $a\in\mathfrak{B}$ ist $\mathfrak{a}_k=\langle z_1,\ldots,z_{k-1},a,z_{k+1},\ldots,z_n\rangle$ aus $\mathfrak{B}$, also $D^k_{\mathfrak{z}}a=D_{\mathfrak{z}}\mathfrak{a}_k=(D_{\mathfrak{x}}\mathfrak{z})^{-1}\cdot D_{\mathfrak{x}}\mathfrak{a}_k\in\mathfrak{B}$ $(1\leqq k\leqq n)$.

Den engen Zusammenhang zwischen „B-uniformisierend" und „B-regulär" zeigt der für die folgenden Ausführungen grundlegende

(3.4) Satz. *Es sei $\Lambda|\mathsf{K}$ separabelerzeugbar; dann gilt:*

$\mathfrak{z}$ ist B-uniformisierend, genau wenn $\mathfrak{z}$ B-regulär und B unverzweigt bezüglich $\mathsf{K}(\mathfrak{z})$ ist. In diesem Falle gilt für jede Differentiation Δ von $\Lambda|\mathsf{K}$:

$$\Delta(\Phi a) = \sum_1^n{}_i \, \Delta(\Phi z_i) \cdot \Phi(D_{\mathfrak{z}}^i a) \qquad (a \in \mathfrak{B}).$$

Beweis. a) Es sei $\mathfrak{z}$ n.g.U. B-regulär (1), B unverzweigt bezüglich $\mathsf{K}(\mathfrak{z})$, P das zu $\Phi_{\mathfrak{z}}$ gehörige Polynom, $\mathfrak{l} = \mathsf{K}(\mathfrak{z}^{(1)})$, $\mathfrak{L}^*$ gemäß (2.5) a) konstruiert (für $j=1$). Nach (2.6) b) ist $\mathfrak{z}$ separierend in $\mathfrak{R}|\mathsf{K}$; jedes $D \in \{D_{\mathfrak{z}}^1, \ldots, D_{\mathfrak{z}}^n\}$ läßt sich wegen $D\mathfrak{l} \subseteq \mathfrak{l}$ (und damit $D\mathfrak{l}_{\mathfrak{R}} \subseteq \mathfrak{l}_{\mathfrak{R}}$) zu einer stetigen Differentiation D^* von $\mathfrak{R}^*$ fortsetzen, und für jedes $a = \sum_0^\infty{}_i \, c_i \cdot (P(\mathfrak{z}))^i \in \mathfrak{B}$ [$c_i \in \mathfrak{L}^*$; (2.5) a)!] gilt nach (2.5) b) $D^* a = \sum_0^\infty{}_i \, D^* c_i \cdot (P(\mathfrak{z}))^i + D(P(\mathfrak{z})) \cdot \sum_0^\infty{}_i \, i \cdot c_i \cdot (P(\mathfrak{z}))^{i-1}$; wegen $D^*\mathfrak{L} \subseteq \mathfrak{L}^* \subseteq \mathfrak{B}^*$ ist $D^* a \in \mathfrak{B}^*$, sogar $D^* a = Da \in \mathfrak{B}^* \cap \mathfrak{R} = \mathfrak{B}$. Also ist $\mathfrak{z}$ B-uniformisierend.

Es sei Δ^i die Differentiation (i) nach $\Phi_{\mathfrak{z}^{(1)}}$ $(2 \leq i \leq n)$ und Δ^1 die triviale Differentiation von $\Lambda|\mathsf{K}$ $(\Delta^1 \Lambda = \{0\})$, ferner D^j die durch $\Phi(D^j c) = \Delta^j(\Phi c)$ $(c \in \mathfrak{L}^*)$ eindeutig bestimmte Differentiation von $\mathfrak{L}^*|\mathsf{K}$ $(1 \leq j \leq n)$. Wegen $\Phi(D_{\mathfrak{z}}^j a) = \Delta^j(\Phi a)$ $(a \in \mathfrak{l})$ stimmt D^j mit $D_{\mathfrak{z}}^j$ in $\mathfrak{l}$ überein; da $\mathfrak{L}^*|\mathfrak{l}$ (ebenso wie $\Lambda|\Phi\mathfrak{l}$) endlich-separabel ist, ist D^j die einzige Differentiation von $\mathfrak{L}^*|\mathsf{K}$ in $\mathfrak{R}^*$ mit dieser Eigenschaft und wird deshalb durch $D_{\mathfrak{z}}^{j*}$ induziert $(1 \leq j \leq n)$, also gilt

$$\Phi(D_{\mathfrak{z}}^{j*} a) = \Phi(D_{\mathfrak{z}}^{j*} c_0) + \Phi(P^{(j)}(\mathfrak{z}) \cdot c_1) = \Delta^j(\Phi a) + P^{(j)}(\Phi_{\mathfrak{z}}) \cdot \Phi c_1 \quad (\Phi a = \Phi c_0!).$$

Für jede Differentiation Δ von $\Lambda|\mathsf{K}$ ist $\Delta(\Phi a) = \sum_1^n{}_j \, \Delta(\Phi z_j) \cdot \Delta^j(\Phi a) = \sum_1^n{}_j \, \Delta(\Phi z_j) \cdot \Phi(D_{\mathfrak{z}}^j a) - \Phi c_1 \cdot \sum_1^n{}_j \, \Delta(\Phi z_j) \cdot P^{(j)}(\Phi_{\mathfrak{z}})$; hiervon ist die letzte Summe gleich Null wegen $P(\Phi_{\mathfrak{z}}) = 0$. — Insbesondere ist

$$\Delta^j(\Phi v_k) = \Delta^j(\Phi z_1) \cdot \Phi(D_{\mathfrak{z}}^1 v_k) + \Phi(D_{\mathfrak{z}}^j v_k) \quad (2 \leq j \leq n,\ 1 \leq k \leq n;\ \mathfrak{v} \text{ aus } \mathfrak{B}).$$

b) Es sei $\mathfrak{y}$ B-uniformisierend, $\mathfrak{z}$ nach (2.1) b) konstruiert. Dann hat die Matrix $\langle \Delta^j(\Phi y_k) \rangle_{2 \leq j \leq n,\, 1 \leq k \leq n}$ den Rang $n-1$; denn anderenfalls wäre für ein $\langle \alpha_2, \ldots, \alpha_n \rangle \neq \langle 0, \ldots, 0 \rangle$ aus Λ $\sum_2^n{}_j \, \alpha_i \cdot \Delta^j(\Phi y_k) = \sum_1^n{}_j \, \alpha_j \cdot \Phi(D_{\mathfrak{z}}^j y_k) = 0$ $\left(\text{mit } \alpha_1 = \sum_2^n{}_j \, \alpha_j \cdot \Delta^j(\Phi z_1);\ 1 \leq k \leq n\right)$, also $|\Phi(D_{\mathfrak{z}}^j y_k)|_{1 \ldots n} = 0$, also $B(D_{\mathfrak{z}}\mathfrak{y}) > 0$ im Widerspruch zu (3.1). Daher ist für ein k $\Phi\mathfrak{y}^{(k)}$ separierend in $\Lambda|\mathsf{K}$, also $\mathfrak{y}$ B-regulär (k). Ist Q das zu $\Phi\mathfrak{y}$ gehörige Polynom, $B(Q(\mathfrak{y})) = m\,(>0)$, $B(u) = 1$, so ist $v = Q(\mathfrak{y}) \cdot u^{-m} \in \mathfrak{B}$, und wegen der Separabilität (k) von Q ist $Q^{(k)}(\mathfrak{y}) = D_{\mathfrak{y}}^k(Q(\mathfrak{y})) = m \cdot u^{m-1} \cdot v \cdot D_{\mathfrak{y}}^k u + u^m \cdot D_{\mathfrak{y}}^k v$ eine Einheit in $\mathfrak{B}$, also $m = 1$, also B unverzweigt bezüglich $\mathsf{K}(\mathfrak{y})$.

§ 4. Gestufte Bewertungen

Als erste Anwendung unserer Ergebnisse übertragen wir den Begriff „uniformisierend" auf gestufte Bewertungen und bringen ihn in Zusammenhang mit den Differentiationen in den zugehörigen allgemeinen Bewertungsringen.

Es sei $\mathfrak{K}|\mathsf{K}$ ein a.F.nV. ($n>1$); $\boldsymbol{B}=\langle B^{\langle 1\rangle},\ldots,B^{\langle n\rangle}\rangle$ heiße eine *gestufte Bewertung* von $\mathfrak{K}|\mathsf{K}$, wenn $B^{\langle i\rangle}$ eine maxdim. Bewertung des a.F.iV. $\mathfrak{K}^{\langle i\rangle}|\mathsf{K}$ ist, wobei $\mathfrak{K}^{\langle n\rangle}=\mathfrak{K}$ und $\mathfrak{K}^{\langle i-1\rangle}=\Lambda^{\langle i\rangle}$ $(1\leq i\leq n)$. Ist $a=a^{\langle n\rangle}\in\mathfrak{B}^{\langle n\rangle},\ldots,a^{\langle i\rangle}\in\mathfrak{B}^{\langle i\rangle}$, so sei $a^{\langle i-1\rangle}=\Phi^{\langle i\rangle}a^{\langle i\rangle}$ $(1\leq i\leq n)$. $\mathfrak{C}_i=\{a\in\mathfrak{K}\,|\,a^{\langle j\rangle}\in\mathfrak{B}^{\langle j\rangle}\ (i\leq j\leq n)\}$ ist ein allgemeiner Bewertungsring von $\mathfrak{K}$ $(1\leq i\leq n)$, $\mathfrak{P}_k=\{a\in\mathfrak{C}_k\,|\,a^{\langle k\rangle}\in\mathfrak{M}^{\langle k\rangle}\}$ ein Primideal von $\mathfrak{C}_i$ $(i\leq k\leq n)$; insbesondere ist $\mathfrak{P}_i$ das Ideal aller Nichteinheiten von $\mathfrak{C}_i$, und es gilt $\mathfrak{B}^{\langle n\rangle}=\mathfrak{C}_n>\cdots>\mathfrak{C}_1>\mathfrak{P}_1>\cdots>\mathfrak{P}_n=\mathfrak{M}^{\langle n\rangle}$ [9]).

Ist $\mathfrak{x}\in\mathfrak{C}_1\times\cdots\times\mathfrak{C}_n$, so bezeichnet $\mathfrak{x}^{\langle i\rangle}$ das i-Tupel $\langle x_1^{\langle i\rangle},\ldots,x_i^{\langle i\rangle}\rangle$ $(0\leq i\leq n)$ und $D^j_{\mathfrak{x}^{\langle i\rangle}}$, falls $\mathfrak{x}^{\langle i\rangle}$ separierend (in $\mathfrak{K}^{\langle i\rangle}|\mathsf{K}$) ist, die Differentiation (j) nach $\mathfrak{x}^{\langle i\rangle}$ von $\mathfrak{K}^{\langle i\rangle}|\mathsf{K}$ $(1\leq j\leq i)$. Als ***B*-uniformisierend** bezeichnen wir jedes $\mathfrak{z}\in\mathfrak{P}_1\times\cdots\times\mathfrak{P}_n$, derart, daß $\mathfrak{z}^{\langle i\rangle}$ $B^{\langle i\rangle}$-uniformisierend (in $\mathfrak{K}^{\langle i\rangle}|\mathsf{K}$) ist $(0\leq i\leq n)$. Für $i=0$ bedeutet dies die Separabilität von $\mathfrak{K}^{\langle 0\rangle}|\mathsf{K}$, die wir im folgenden stets voraussetzen; aus ihr folgt, daß $\mathfrak{K}^{\langle j\rangle}|\mathsf{K}$ separabelerzeugbar ist [$0\leq j\leq n$; (2.6) b)!].

(4.1) *$\mathfrak{z}$ ist **B**-uniformisierend, genau wenn $B^{\langle i\rangle}(z_i^{\langle i\rangle})=1$ $(1\leq i\leq n)$ ist.*

Beweis. Es sei $B^{\langle i\rangle}(z_i^{\langle i\rangle})=1$ $(1\leq i\leq n)$. Ist $\mathfrak{z}^{\langle j-1\rangle}$ $B^{\langle j-1\rangle}$-uniformisierend (für $j=1$ trivial!), so ist nach (2.1) $\mathfrak{z}^{\langle j\rangle}$ $B^{\langle j\rangle}$-regulär und $B^{\langle j\rangle}$ unverzweigt bezüglich $\mathsf{K}(\mathfrak{z}^{\langle j\rangle})$, also ist $\mathfrak{z}^{\langle j\rangle}$ nach (3.4) $B^{\langle j\rangle}$-uniformisierend. — Es sei $\mathfrak{z}$ ***B*-uniformisierend**; nach (3.4) ist $\mathfrak{z}^{\langle j\rangle}$ $B^{\langle j\rangle}$-regulär und $B^{\langle j\rangle}$ unverzweigt bezüglich $\mathsf{K}(\mathfrak{z}^{\langle j\rangle})$, also nach (2.2) (wegen $z_j^{\langle j-1\rangle}=0$) $B^{\langle j\rangle}(z_j^{\langle j\rangle})=1$ $(1\leq j\leq n)$.

In $\mathfrak{P}_1\times\cdots\times\mathfrak{P}_n$ sind die ***B*-uniformisierenden** n-Tupel nicht nur ausgezeichnet in bezug auf die Differentiationen in den Bewertungsringen $\mathfrak{B}^{\langle i\rangle}$, sondern auch hinsichtlich der Differentiationen in den allgemeinen Bewertungsringen $\mathfrak{C}_i<\mathfrak{K}$ $(1\leq i\leq n)$, wie der folgende Satz zeigt:

(4.2) Satz. *Ein separierendes $\mathfrak{z}\in\mathfrak{P}_1\times\cdots\times\mathfrak{P}_n$ ist **B**-uniformisierend, genau wenn $D^j_{\mathfrak{z}}\mathfrak{C}_k\subseteq\mathfrak{C}_k$ $(1\leq j\leq k\leq n)$ gilt.*

Wir beweisen zunächst:

(4.3) *Es sei $\mathfrak{z}\in\mathfrak{P}_1\times\cdots\times\mathfrak{P}_n$, für ein $i_0\in\{1,\ldots,n-1\}$ sei $\mathfrak{z}^{\langle i_0\rangle}$ separierend und $\mathfrak{z}^{\langle i\rangle}$ $B^{\langle i\rangle}$-uniformisierend $(i_0<i\leq n)$; dann ist*

$$D^j_{\mathfrak{z}^{\langle i_0\rangle}}a^{\langle i_0\rangle}=(D^j_{\mathfrak{z}}a)^{\langle i_0\rangle}\quad(a\in\mathfrak{C}_{i_0+1},\ 1\leq j\leq i_0).$$

Beweise. (4.3): Gilt für ein $i>i_0$ $D^j_{\mathfrak{z}^{\langle i\rangle}}a^{\langle i\rangle}=(D^j_{\mathfrak{z}}a)^{\langle i\rangle}$ $(a\in\mathfrak{C}_{i_0+1},\ 1\leq j\leq i_0$; für $i=n$ trivial!), so ist nach (3.4) $D^j_{\mathfrak{z}^{\langle i-1\rangle}}a^{\langle i-1\rangle}=\sum_1^i{}_k\,D^j_{\mathfrak{z}^{\langle i-1\rangle}}z_k^{\langle i-1\rangle}\cdot(D^k_{\mathfrak{z}^{\langle i\rangle}}a^{\langle i\rangle})^{\langle i-1\rangle}=$ (wegen $z_i^{\langle i-1\rangle}=0$) $=(D^j_{\mathfrak{z}^{\langle i\rangle}}a^{\langle i\rangle})^{\langle i-1\rangle}=(D^j_{\mathfrak{z}}a)^{\langle i-1\rangle}$.

(4.2): Es sei $\mathfrak{z}$ ***B*-uniformisierend**; dann ist für alle $a\in\mathfrak{C}_i$ $(D^j_{\mathfrak{z}}a)^{\langle i\rangle}=[(4.3)!]=D^j_{\mathfrak{z}^{\langle i\rangle}}a^{\langle i\rangle}\in\mathfrak{B}^{\langle i\rangle}$, also $D^j_{\mathfrak{z}}a\in\mathfrak{C}_i$ $(1\leq j\leq i\leq n)$. — Es sei $D^j_{\mathfrak{z}}\mathfrak{C}_k\subseteq\mathfrak{C}_k$ $(1\leq j\leq k\leq n)$. Ist $\mathfrak{z}^{\langle i\rangle}$ $B^{\langle i\rangle}$-uniformisierend (dies trifft zu für $i=n$!), so ist nach (3.4) $\mathfrak{z}^{\langle i\rangle}$ $B^{\langle i\rangle}$-regulär (i), $\mathfrak{z}^{\langle i-1\rangle}$ separierend. Ist $\alpha\in\mathfrak{B}^{\langle i-1\rangle}$, $a\in\mathfrak{C}_{i-1}$ ein Repräsentant von α, so ist $D^j_{\mathfrak{z}}a\in\mathfrak{C}_{i-1}$, also $D^j_{\mathfrak{z}^{\langle i-1\rangle}}\alpha=[(4.3)!]=(D^j_{\mathfrak{z}}a)^{\langle i-1\rangle}\in\mathfrak{B}^{\langle i-1\rangle}$ $(1\leq j\leq i-1)$ also $\mathfrak{z}^{\langle i-1\rangle}$ $B^{\langle i-1\rangle}$-uniformisierend.

[9]) Vgl. W. KRULL: Allgemeine Bewertungstheorie. J. f. Math. **167**, 160—196 (1931), insb. § 5.

Für $\boldsymbol{B}$-uniformisierendes $\mathfrak{z}$ wird $\mathfrak{C}_k$ durch $D_{\mathfrak{z}}^1, \ldots, D_{\mathfrak{z}}^k$ in sich abgebildet, dagegen ist $D_{\mathfrak{z}}^j \mathfrak{C}_k \not\subseteq \mathfrak{C}_k$ für $j > k$ [z.B. ist $z_j \cdot z_k^{-1} \in \mathfrak{P}_j < \mathfrak{C}_k$, aber $D_{\mathfrak{z}}^j (z_j \cdot z_k^{-1}) = z_k^{-1} \notin \mathfrak{C}_k$]; also ist $D_{\mathfrak{z}} \mathfrak{C}_k \subseteq \mathfrak{C}_k$ nur für $k = n$ erfüllt. Dies kann jedoch nicht als Mangel unserer Definition „$\boldsymbol{B}$-uniformisierend" angesehen werden, denn es gibt überhaupt kein $\mathfrak{y}$ aus $\mathfrak{C}_n$ mit $D_{\mathfrak{y}} \mathfrak{C}_k \subseteq \mathfrak{C}_k$ $(1 \leq k \leq n)$; es gilt sogar:

(4.4) *Für kein* $\mathfrak{y}$ *aus* $\mathfrak{C}_n$ *gilt zugleich* $D_{\mathfrak{y}} \mathfrak{C}_n \subseteq \mathfrak{C}_n$ *und* $D_{\mathfrak{y}} \mathfrak{C}_{n-1} \subseteq \mathfrak{C}_{n-1}$.

Beweis. Ist $D_{\mathfrak{y}} \mathfrak{C}_n \subseteq \mathfrak{C}_n$, also $\mathfrak{y}$ $B^{\langle n \rangle}$-uniformisierend, $u \in \mathfrak{P}_n$, $v \in \mathfrak{C}_n - \mathfrak{C}_{n-1}$, so ist $u \cdot v^k \in \mathfrak{P}_n \subseteq \mathfrak{C}_{n-1}$ und $\left(D_{\mathfrak{y}}(u \cdot v^k)\right)^{\langle n-1 \rangle} = (v^{\langle n-1 \rangle})^k \cdot (D_{\mathfrak{y}}^i u)^{\langle n-1 \rangle}$ $(1 \leq i \leq n)$. Nach (3.4) ist $\mathfrak{y}$ $B^{\langle n \rangle}$-regulär, also, wenn P das zu $\Phi^{\langle n \rangle} \mathfrak{y}$ gehörige Polynom ist, $D_{\mathfrak{y}}^i \left(P(\mathfrak{y})\right) = P^{(i)}(\mathfrak{y}) \notin \mathfrak{P}_n$ für ein j. Für $u = P(\mathfrak{y})$ ist daher $(D_{\mathfrak{y}}^j u)^{\langle n-1 \rangle} \neq 0$, also, wegen $B^{\langle n-1 \rangle}(v^{\langle n-1 \rangle}) < 0$, für genügend großes k $D_{\mathfrak{y}}^j (u \cdot v^k) \notin \mathfrak{C}_{n-1}$.

Wir beweisen noch, in Analogie zu (3.1):

(4.5) *Es sei* $\mathfrak{z}$ $\boldsymbol{B}$*-uniformisierend,* $\mathfrak{y} \in \mathfrak{P}_1 \times \cdots \times \mathfrak{P}_n$. *Dann ist* $|D_{\mathfrak{z}}^j y_i|_{1 \ldots k} \in \mathfrak{C}_k$ $(1 \leq k \leq n)$; $\mathfrak{y}$ *ist* $\boldsymbol{B}$*-uniformisierend, genau wenn* $|D_{\mathfrak{z}}^j v_i|_{1 \ldots k} \notin \mathfrak{P}_k$ $(1 \leq k \leq n)$ *ist.*

Beweis. $|D_{\mathfrak{z}}^j y_i|_{1 \ldots k} \in \mathfrak{C}_k$ folgt aus (4.2). — Nach (4.3) ist $(|D_{\mathfrak{z}}^j y_i|_{1 \ldots k})^{\langle k \rangle} = |D_{\mathfrak{z}^{\langle k \rangle}}^j y_i^{\langle k \rangle}|_{1 \ldots k}$; also ist $|D_{\mathfrak{z}}^j y_i|_{1 \ldots k} \notin \mathfrak{P}_k$, genau wenn $|D_{\mathfrak{z}^{\langle k \rangle}}^j y_i^{\langle k \rangle}|_{1 \ldots k} \notin \mathfrak{M}^{\langle k \rangle}$, genau wenn $\mathfrak{y}^{\langle k \rangle}$ $B^{\langle k \rangle}$-uniformisierend ist $(1 \leq k \leq n)$.

§ 5. Differenten und Differentiale

Es sei B eine maxdim. Bewertung des a.F. nV. $\mathfrak{R} | \mathsf{K}$, $\mathsf{K} \subseteq \mathfrak{k} \subseteq \mathfrak{R}$, $\mathfrak{R} | \mathfrak{k}$ endlich-algebraisch. Dann definieren wir nach Ch. IV § 8 die *B-Differente* von $\mathfrak{R} | \mathfrak{k}$ („differential exponent"): $\mathrm{Diff}_{\mathfrak{R}|\mathfrak{k}} B = \max \{m \mid$ aus $B(a) \geq -m$ folgt $\mathrm{Sp}_{\mathfrak{R}^* | \mathfrak{k}^*} a \in \mathfrak{b}^*$ $(a \in \mathfrak{R})\}$.

Aus Ch. IV § 8, Theorem 7 [10]) bzw. 8 folgt mit Hilfe von (2.3) der „Exponentensatz":

(5.1) a) $\mathrm{Diff}_{\mathfrak{R}|\mathfrak{k}} B < \infty$, *genau wenn* $\mathfrak{R} | \mathfrak{k}$ *separabel ist;*

b) $\mathrm{Diff}_{\mathfrak{R}|\mathfrak{k}} B \geq e - 1$; *hierbei gilt das Gleichheitszeichen, genau wenn* B *separabel bezüglich* $\mathfrak{k}$ *und* e *durch* $Ch(\mathfrak{R})$ *unteilbar ist;*

sowie der „Schachtelungssatz":

(5.2) *Ist* $\mathfrak{l} \subseteq \mathfrak{k} \subseteq \mathfrak{R}$, $\mathfrak{R} | \mathfrak{l}$ *endlich-algebraisch, so ist*

$$\mathrm{Diff}_{\mathfrak{R}|\mathfrak{l}} B = \mathrm{Diff}_{\mathfrak{R}|\mathfrak{k}} B + e_{\mathfrak{R}|\mathfrak{k}} \cdot \mathrm{Diff}_{\mathfrak{k}|\mathfrak{l}} b.$$

Als Differentiale eines a.F. 1V. $\mathfrak{R} | \mathsf{K}$ definiert Ch. gewisse Linearfunktionen auf dem $\mathsf{K}_{\mathfrak{R}}$-Modul aller „repartitions" und überträgt auf sie in einer natürlichen Weise die Bewertungen von $\mathfrak{R} | \mathsf{K}$ (Ch. II §§ 4—6). Ist $\mathfrak{R} | \mathsf{K}$ separabel-erzeugbar, so läßt sich jedes Differential in der Form $u \cdot dx$ $(u, x \in \mathfrak{R})$ darstellen (Ch. VI § 3). Differente und Differentiale hängen durch die „Riemannsche Formel"

(5.3) $B(dx) = \mathrm{Diff}_{\mathfrak{R}|\mathsf{K}(x)} B \qquad (x \in \mathfrak{B},$ *separierend in* $\mathfrak{R} | \mathsf{K})$

[10]) Ch. behandelt nur den separablen Fall. Ist $\mathfrak{R} | \mathfrak{k}$ inseparabel, so gilt dasselbe für $\mathfrak{R}^* | \mathfrak{k}^*$ (Ch. IV S. 68 oben), also ist $Sp_{\mathfrak{R}^*|\mathfrak{k}^*} \mathfrak{R} = \{0\}$, also $\mathrm{Diff}_{\mathfrak{R}|\mathfrak{k}} B = \infty$; und nach (2.6) a) ist stets B inseparabel bez. $\mathfrak{k}$ oder e durch $Ch(\mathfrak{R})$ teilbar.

zusammen. Sie wird in CH. VI § 3, Th. 7 nur für den Fall $\mathsf{K} = \mathsf{K}_{\mathfrak{R}}$ bewiesen, gilt jedoch auch ohne diese Voraussetzung, da für endlich-separables $\tilde{\mathsf{K}}|\mathsf{K}$ stets $\mathrm{Diff}_{\tilde{\mathsf{K}}(x)|\mathsf{K}(x)}\tilde{b} = 0$ [11]) ist.

Im folgenden konstruieren wir für separabelerzeugbare a.F. nV. $\mathfrak{R}|\mathsf{K}$ Differentiale (n-ten Grades), übertragen auf sie die maxdim. Bewertungen von $\mathfrak{R}|\mathsf{K}$ derart, daß im Falle $n = 1$ Übereinstimmung mit CH. besteht, und beweisen schließlich eine Verallgemeinerung der Riemannschen Formel.

Als *Differentiale* von $\mathfrak{R}|\mathsf{K}$ bezeichnen wir die Elemente $u \cdot d\mathfrak{x}\,(= u \cdot d_{\mathfrak{R}|\mathsf{K}}\mathfrak{x}) = u \cdot \partial\mathfrak{x} + H_{\mathfrak{R}|\mathsf{K}}$ des Restklassenmoduls $\boldsymbol{d}_{\mathfrak{R}|\mathsf{K}} = \mathfrak{R} \cdot \partial\mathfrak{R}^n / H_{\mathfrak{R}|\mathsf{K}}$, wobei ∂ eine eineindeutige Abbildung von $\mathfrak{R}^n = \mathfrak{R} \times \cdots \times \mathfrak{R}$, $\mathfrak{R} \cdot \partial\mathfrak{R}^n$ den durch $\partial\mathfrak{R}^n$ erzeugten freien $\mathfrak{R}$-Modul und $H_{\mathfrak{R}|\mathsf{K}}$ den durch $\{\partial\mathfrak{v} - D_{\mathfrak{z}}\mathfrak{v} \cdot \partial\mathfrak{z} \mid \mathfrak{v}, \mathfrak{z} \in \mathfrak{R}^n,\ \mathfrak{z}$ separierend in $\mathfrak{R}|\mathsf{K}\}$ erzeugten Untermodul bezeichnet. Ersichtlich ist $H_{\mathfrak{R}|\mathsf{K}} = H_{\mathfrak{R}|\mathsf{K}_{\mathfrak{R}}}$, also auch $\boldsymbol{d}_{\mathfrak{R}|\mathsf{K}} = \boldsymbol{d}_{\mathfrak{R}|\mathsf{K}_{\mathfrak{R}}}$, und es gilt:

(5.4) $\boldsymbol{d}_{\mathfrak{R}|\mathsf{K}}$ *ist ein eingliedriger (unitärer)* $\mathfrak{R}$*-Modul; für separierendes* $\mathfrak{z}$ *aus* $\mathfrak{R}$ *ist* $d\mathfrak{z} \neq 0$ *und* $d\mathfrak{v} = D_{\mathfrak{z}}\mathfrak{v} \cdot d\mathfrak{z}$ ($\mathfrak{v}$ aus $\mathfrak{R}$).

$\boldsymbol{d}_{\mathfrak{R}|\mathsf{K}}$ ist durch (5.4) im wesentlichen eindeutig bestimmt; insbesondere gilt (5.4) für die Chevalleyschen Differentiale (CH. II § 6, VI § 3 Th. 4 und § 5 Th. 9), die wir deshalb mit den Elementen $\in \boldsymbol{d}_{\mathfrak{R}|\mathsf{K}}$ im Falle $n = 1$ identifizieren können.

Für die Übertragung der Bewertungen auf die Differentiale brauchen wir eine Verallgemeinerung der Differente auf a.F. nV. $\mathfrak{R}|\mathsf{K}$: $\mathrm{Diff}_{\mathfrak{R}|\mathsf{K}}B = \min\{\mathrm{Diff}_{\mathfrak{R}|\mathsf{K}(\mathfrak{x})}B \mid \mathfrak{x}$ Tr.-B. von $\mathfrak{R}|\mathsf{K}\}$. Ersichtlich ist $\mathrm{Diff}_{\mathfrak{R}|\mathsf{K}}B < \infty$, genau wenn $\mathfrak{R}|\mathsf{K}$ separabelerzeugbar ist, und aus (5.1) b) und (2.1) folgt unmittelbar:

(5.5) $\mathrm{Diff}_{\mathfrak{R}|\mathsf{K}}B = 0$, *genau wenn für ein in* $\mathfrak{R}|\mathsf{K}$ *separierendes* $\mathfrak{x}$ B *separabel und unverzweigt bezüglich* $\mathsf{K}(\mathfrak{x})$ *ist, insbesondere wenn* $\Lambda|\mathsf{K}$ *separabelerzeugbar ist.*

Bei vollkommenem K ist also stets $\mathrm{Diff}_{\mathfrak{R}|\mathsf{K}}B = 0$. Daß dies für separabelerzeugbare a.F. nV. (schon im Falle $n = 1$) nicht allgemein gilt, zeigt folgendes Beispiel [12]): Es sei $Ch(\mathsf{K}_0) = p > 0$, $\mathsf{K} = \mathsf{K}_0(\xi, \eta)$, $\mathfrak{k} = \mathsf{K}(x)$, $\mathfrak{R} = \mathfrak{k}(y)$, wobei $\langle x, \xi, \eta\rangle$ alg.-unabh. bezüglich K_0 und y Nullstelle des bezüglich $\mathfrak{k}$ irreduziblen Polynoms $Y^p - \eta \cdot (x^p - \xi) - \eta$ sei; B liege über der Bewertung $b_{x, x^p - \xi}$ von $\mathfrak{k}|\mathsf{K}$. Dann ist $\Phi x = \sqrt[p]{\xi} \in \Lambda$, $\Phi y = \sqrt[p]{\eta} \in \Lambda$ und x separierend in $\mathfrak{R}|\mathsf{K}$. Da $\mathsf{K}(\sqrt{\xi}, \sqrt[p]{\eta})|\mathsf{K}$ nichteinfach ist, gilt dasselbe für $\Lambda|\mathsf{K}$, und für kein $\zeta \in \Lambda$ ist $\Lambda|\mathsf{K}(\zeta)$ separabel [13]), also ist für jedes $z \in \mathfrak{R} - \mathsf{K}_{\mathfrak{R}}$ B inseparabel bezüglich $\mathsf{K}(z)$, also nach (5.1) b) $\mathrm{Diff}_{\mathfrak{R}|\mathsf{K}(z)}B > 0$.

Wir übertragen nun die maxdim. Bewertungen B von $\mathfrak{R}|\mathsf{K}$ auf $\boldsymbol{d}_{\mathfrak{R}|\mathsf{K}}$ durch die Festsetzung: $B(u \cdot d\mathfrak{z}) = B(u) + B(d\mathfrak{z})$ und $B(d\mathfrak{z}) = \mathrm{Diff}_{\mathfrak{R}|\mathsf{K}}B$ für B-uniformisierendes $\mathfrak{z}$, $u \in \mathfrak{R}$ [(3.1), (5.4)!]. Im Falle $n = 1$ stimmt dies mit CH.

[11]) Ist $b_{z,P}$ die unter $\tilde{b} = \tilde{b}_{z,\tilde{P}}$ ($z = x$ oder $= x^{-1}$) liegende Bewertung von $\mathsf{K}(x)$, so ist P durch $\tilde{P}$, aber nicht durch $\tilde{P}^2$ teilbar, also $\tilde{b}$ unverzweigt bezüglich $\mathfrak{k}$. Ferner ist $\tilde{b}$ separabel bezüglich $\mathfrak{k}$, also gilt die Gleichung nach (5.1) b).

[12]) $\mathrm{Diff}_{\mathfrak{R}|\mathsf{K}}B > 0$ ist jedoch als Ausnahmefall anzusehen, denn jedes projektive Bewertungssystem $\mathfrak{S}_{|\mathfrak{z}|}$ ($\mathfrak{z}$ separierend in $\mathfrak{R}|\mathsf{K}$) enthält höchstens endlich viele Bewertungen B dieser Art [(6.6*)!].

[13]) Vgl. G. PICKERT: Einführung in die Höhere Algebra, Göttingen 1951; insbesondere § 31 S. 196 unten und Aufgabe 2.

überein; denn es gibt ein separierendes $x \in \mathfrak{B}$ mit $\mathrm{Diff}_{\mathfrak{R}|\mathsf{K}(x)} B = \mathrm{Diff}_{\mathfrak{R}|\mathsf{K}} B$, nach (5.3) ist $B(dx) = \mathrm{Diff}_{\mathfrak{R}|\mathsf{K}} B$ und x B-uniformisierend wegen $B(dx) = \mathrm{Diff}_{\mathfrak{R}|\mathsf{K}} B \leq \mathrm{Diff}_{\mathfrak{R}|\mathsf{K}(v)} B = B(dv) = B(D_x v) + B(dx)$ ($v \in \mathfrak{B}$, separierend). — Aus (3.1) folgt unmittelbar:

(5.6) $B(d\mathfrak{v}) \geq \mathrm{Diff}_{\mathfrak{R}|\mathsf{K}} B$ ($\mathfrak{v}$ *aus* $\mathfrak{B}$); *das Gleichheitszeichen gilt, genau wenn* $\mathfrak{v}$ *B-uniformisierend ist.*

Um die Riemannsche Formel auf Differentiale $d\mathfrak{x}$ B-zulässiger[14]) separierender n-Tupel $\mathfrak{x}$ eines a.F.nV. $\mathfrak{R}|\mathsf{K}$ verallgemeinern zu können, brauchen wir den Begriff der „Semiregularität": ein B-zulässiges (k) $\mathfrak{x}$ heiße *B-semiregulär* (k), wenn für ein $z \in \mathfrak{R}\ \langle x_1, \ldots, x_{k-1}, z, x_{k+1}, \ldots, x_n \rangle$ B-uniformisierend ist.

(5.7) *Jedes B-reguläre (k) $\mathfrak{x}$ ist B-semiregulär (k).*

(5.8) *Ist $\Lambda|\mathsf{K}$ separabelerzeugbar und $\mathfrak{x}$ B-semiregulär (k), so ist* $\mathrm{Diff}_{\mathfrak{R}|\mathsf{K}(\mathfrak{x}^{(k)})} B = 0$.

Beweise[15]). (5.7): Für $B(z) = 1$ ist nach (2.1) $\mathfrak{z} = \langle z, x_2, \ldots, x_n \rangle$ B-regulär (1) und B unverzweigt bezüglich $\mathsf{K}(\mathfrak{z})$; nach (3.4) ist $\mathfrak{z}$ B-uniformisierend.

(5.8): Für ein $z \in \mathfrak{B}$ ist $\mathfrak{z} = \langle z, x_2, \ldots, x_n \rangle$ B-uniformisierend, nach (3.4) ist B separabel und unverzweigt bezüglich $\mathsf{K}(\mathfrak{z})$. B ist eine Bewertung von $\mathfrak{R}|\mathsf{K}(\mathfrak{x}^{(1)})$, z separierend in $\mathfrak{R}|\mathsf{K}(\mathfrak{x}^{(1)})$, also ist nach (5.5) $\mathrm{Diff}_{\mathfrak{R}|\mathsf{K}(\mathfrak{x}^{(1)})} B = 0$.

Daß es tatsächlich a.F.nV. $\mathfrak{R}|\mathsf{K}$ (auch solche mit vollkommenem Koeffizientenkörper K) gibt, die separierende B-zulässige (k) aber B-nichtsemireguläre (k) $\mathfrak{x}$ besitzen, zeigt das obenstehende Beispiel: Ersichtlich ist $\langle x, \xi, \eta \rangle$ separierend in $\mathfrak{R}|\mathsf{K}_0$ und B-zulässig (1); wählt man K_0 vollkommen, so ist wegen $\mathrm{Diff}_{\mathfrak{R}|\mathsf{K}} B > 0$ nach (5.8) $\langle x, \xi, \eta \rangle$ B-nichtsemiregulär (1) (in $\mathfrak{R}|\mathsf{K}_0$).

Da im Falle $n = 1$ ersichtlich jedes B-zulässige x (d.h. $x \in \mathfrak{B} - \mathsf{K}_{\mathfrak{R}}$) B-semiregulär ist, ist der folgende Satz eine Verallgemeinerung der Riemannschen Formel:

(5.9) Satz. *Ist $\mathfrak{x}$ B-zulässig (k) und separierend in $\mathfrak{R}|\mathsf{K}$, so ist*

$$B(d\mathfrak{x}) \geq \mathrm{Diff}_{\mathfrak{R}|\mathsf{K}(\mathfrak{x})} B - \mathrm{Diff}_{\mathfrak{R}|\mathsf{K}(\mathfrak{x}^{(k)})} B - \mathrm{Diff}_{\mathfrak{R}|\mathsf{K}} B;$$

das Gleichheitszeichen gilt, genau wenn $\mathfrak{x}$ B-semiregulär (k) ist.

Ist $\Lambda|\mathsf{K}$ separabelerzeugbar, so gilt für jedes B-semireguläre $\mathfrak{x}$ sogar

$$B(d\mathfrak{x}) = \mathrm{Diff}_{\mathfrak{R}|\mathsf{K}(\mathfrak{x})} B.$$

Beweis[15]). Es sei $\mathfrak{l} = \mathsf{K}(\mathfrak{x}^{(1)})$, $z \in \mathfrak{B}$, $\mathfrak{z} = \langle z, x_2, \ldots, x_n \rangle$; dann ist $d\mathfrak{z} = D_{\mathfrak{x}} \mathfrak{z} \cdot d\mathfrak{x} = D^1_{\mathfrak{x}} z \cdot d\mathfrak{x}$ und $d_{\mathfrak{R}|\mathfrak{l}} z = D^1_{\mathfrak{x}} z \cdot d_{\mathfrak{R}|\mathfrak{l}} x_1$. Nach (5.3) ist $B(d_{\mathfrak{R}|\mathfrak{l}} x_1) = \mathrm{Diff}_{\mathfrak{R}|\mathsf{K}(\mathfrak{x})} B$. Wählt man z B-uniformisierend in $\mathfrak{R}|\mathfrak{l}$, so ist nach (5.6) $B(d_{\mathfrak{R}|\mathfrak{l}} z) = \mathrm{Diff}_{\mathfrak{R}|\mathfrak{l}} B$ ($< \infty$) und $B(d\mathfrak{z}) \geq \mathrm{Diff}_{\mathfrak{R}|\mathsf{K}} B$, also $B(d\mathfrak{x}) \geq \mathrm{Diff}_{\mathfrak{R}|\mathsf{K}} B - \mathrm{Diff}_{\mathfrak{R}|\mathfrak{l}} B + \mathrm{Diff}_{\mathfrak{R}|\mathsf{K}(\mathfrak{x})} B$; dabei gelten die Ungleichheitszeichen, wenn $\mathfrak{x}$ B-nichtsemiregulär (1) ist. Ist $\mathfrak{x}$ B-semiregulär (1), so gibt es ein $z \in \mathfrak{B}$ derart, daß $\mathfrak{z}$ B-uniformisierend ist, also gilt wegen (5.6) die letztere Ungleichung mit $\leq$ statt $\geq$. Der Rest folgt aus (5.5) und (5.8).

[14]) Für B-nichtzulässige $\mathfrak{v}$ ist ein einfacher Zusammenhang zwischen $B(d\mathfrak{v})$ und $\mathrm{Diff}_{\mathfrak{R}|\mathsf{K}(\mathfrak{v})} B$ nicht zu erwarten. Ist z.B. $n > 1$, $\mathfrak{x}$ B-zulässig, $B(x_n) > 0$, $\mathfrak{v}_i = \langle x_1 \cdot x_n^i, x_2, \ldots, x_n \rangle$ (B-nichtzulässig für $i \neq 0$!), so sind die $B(d\mathfrak{v}_i)$ paarweise verschieden, dagegen [wegen $\mathsf{K}(\mathfrak{v}_i) = \mathsf{K}(\mathfrak{x})$] alle B-Differenten $\mathrm{Diff}_{\mathfrak{R}|\mathsf{K}(\mathfrak{v}_i)} B$ gleich.

[15]) Im Beweis wird o.B.d.A. $k = 1$ angenommen.

Als eine befriedigende Verallgemeinerung der Riemannschen Formel kann dieser Satz nur im Falle $Ch(\mathfrak{R}) = 0$ angesehen werden, in dem stets $B(d\mathfrak{x}) = \mathrm{Diff}_{\mathfrak{R}|\mathsf{K}(\mathfrak{x})} B$ ($\mathfrak{x}$ B-zulässig) gilt[4]). Bei beliebiger Charakteristik und vollkommenem Koeffizientenkörper gilt, wie in § 6 gezeigt wird, diese Gleichung wenigstens für fast alle Bewertungen B des zu $\mathfrak{x}$ gehörigen projektiven Bewertungssystems ($\mathfrak{x}$ separierend).

§ 6. Über projektive Bewertungssysteme

Mit Hilfe der in §§ 2–3 bewiesenen Tatsachen wollen wir nun die Ergebnisse von KR. § 6 über projektive Bewertungssysteme durch einige Sätze ergänzen.

Zunächst beweisen wir einige Hilfssätze über die „projektiven Transformationen" (von $\mathfrak{R}|\mathsf{K}$), d.h. die durch

$$\Omega\mathfrak{x} = \left\langle \frac{\sum_0^n{}_i\, \omega_{1i} x_i}{\sum_0^n{}_i\, \omega_{0i} x_i}, \ldots, \frac{\sum_0^n{}_i\, \omega_{ni} x_i}{\sum_0^n{}_i\, \omega_{0i} x_i} \right\rangle \quad (\text{mit } x_0 = 1;\ \langle \omega_{ij} \rangle_{0\ldots n} \text{ regulär aus } \mathsf{K})$$

definierten eineindeutigen Abbildungen Ω von der Menge $\mathfrak{X}$ aller Tr.-B. $\mathfrak{x}$ von $\mathfrak{R}|\mathsf{K}$ auf $\mathfrak{X}$. Ihre Gesamtheit bildet eine Gruppe $P_{\mathfrak{R}|\mathsf{K}}$, echte Obergruppe der Gruppe $G_{\mathfrak{R}|\mathsf{K}}$ aller ganzlinearen Transformationen Γ:

$$\Gamma\mathfrak{x} = \left\langle \sum_0^n{}_i\, \gamma_{1i} x_i, \ldots, \sum_0^n{}_i\, \gamma_{ni} x_i \right\rangle \quad (\text{mit } x_0 = 1;\ \langle \gamma_{ij} \rangle_{1\ldots n} \text{ regulär aus } \mathsf{K}).$$

Mit L_Ω bezeichnen wir das (bis auf einen Faktor $\in \mathsf{K}^*$ eindeutig bestimmte) Polynom $\omega_{00} + \omega_{01} X_1 + \cdots + \omega_{0n} X_n$. Es gilt

(6.1) a) $L_\Omega(\mathfrak{x}) \cdot L_{\Omega^{-1}}(\Omega\mathfrak{x}) \in \mathsf{K}^* \qquad (\Omega \in P_{\mathfrak{R}|\mathsf{K}},\ \mathfrak{x} \in \mathfrak{X})$.

b) *Jedes* $\Omega \in P_{\mathfrak{R}|\mathsf{K}} - G_{\mathfrak{R}|\mathsf{K}}$ *läßt sich als Produkt* $\Omega = \widetilde{\Gamma} \cdot \mathsf{T} \cdot \Gamma$ *für geeignete* $\Gamma, \widetilde{\Gamma} \in G_{\mathfrak{R}|\mathsf{K}}$ *mit* $L_\Omega = \gamma_{n0} + \gamma_{n1} X_1 + \cdots + \gamma_{nn} X_n$ *darstellen* ($\mathsf{T} \in P_{\mathfrak{R}|\mathsf{K}}$ *sei durch* $\mathsf{T}\mathfrak{x} = \left\langle \frac{x_1}{x_n}, \ldots, \frac{x_{n-1}}{x_n}, \frac{1}{x_n} \right\rangle$ $(\mathfrak{x} \in \mathfrak{X})$ *definiert*).

c) $D_{\mathfrak{x}}(\Omega\mathfrak{x}) \cdot \left(L_\Omega(\mathfrak{x})\right)^{n+1} \in \mathsf{K}^* \qquad (\Omega \in P_{\mathfrak{R}|\mathsf{K}},\ \mathfrak{x} \in \mathfrak{X})$.

Beweis. a) Es sei $\mathfrak{y} = \Omega\mathfrak{x}$, $\sum_0^n{}_k\, \widetilde{\omega}_{ik} \cdot \omega_{kj} = \delta_{ij}$ $(i, j = 0, \ldots, n)$; dann ist

$$L_{\Omega^{-1}}(\mathfrak{y}) = \sum_0^n{}_k\, \widetilde{\omega}_{0k} \cdot y_k = \left(\sum_{0}^{n}{}_{k,j} \widetilde{\omega}_{0k} \cdot \omega_{kj} \cdot x_j\right) \cdot \left(\sum_0^n{}_k\, \omega_{0k} \cdot x_k\right)^{-1} = \left(L_\Omega(\mathfrak{x})\right)^{-1} \ (x_0 = y_0 = 1).$$

b) Wegen $\Omega \notin G_{\mathfrak{R}|\mathsf{K}}$ ist $\omega_{0k} \neq 0$ für ein $k \in \{1, \ldots, n\}$. Ist $\Gamma\mathfrak{x} = \left\langle x_1, \ldots, x_{k-1}, x_n, x_{k+1}, \ldots, x_{n-1}, \sum_0^n{}_i\, \omega_{0i} \cdot x_i \right\rangle$ $(\mathfrak{x} \in \mathfrak{X},\ x_0 = 1)$, so ist $L_\Omega = \gamma_{n0} + \gamma_{n1} X_1 + \cdots + \gamma_{nn} X_n$ und, wie man leicht nachrechnet, $\widetilde{\Gamma} = \Omega \cdot \Gamma^{-1} \cdot \mathsf{T}^{-1} \in G_{\mathfrak{R}|\mathsf{K}}$.

c) Es sei $\Omega = \widetilde{\Gamma} \cdot \mathsf{T} \cdot \Gamma$ gemäß b), $\mathfrak{u} = \Gamma\mathfrak{x}$, $\mathfrak{v} = \mathsf{T}\mathfrak{u}$, $\mathfrak{y} = \widetilde{\Gamma}\mathfrak{v}$. Dann ist $u_n = \sum_0^n{}_i\, \gamma_{ni} \cdot x_i = L_\Omega(\mathfrak{x})$, $D_{\mathfrak{x}}(\Omega\mathfrak{x}) = D_{\mathfrak{x}}\mathfrak{u} \cdot D_{\mathfrak{u}}\mathfrak{v} \cdot D_{\mathfrak{v}}\mathfrak{y} = \alpha \cdot D_{\mathfrak{u}}\mathfrak{v} = -\alpha \cdot u_n^{-(n+1)}$ für ein $\alpha \in \mathsf{K}^*$.

Es sei $\mathfrak{x}$ eine Tr.-B. des a. F. nV. $\mathfrak{R}|\mathsf{K}$; $\mathfrak{s}_\mathfrak{x}$ bezeichne die Menge aller Polynombewertungen $b_{\mathfrak{x},P}$ von $\mathsf{K}(\mathfrak{x})|\mathsf{K}$, $\mathfrak{S}_\mathfrak{x}$ das über $\mathfrak{s}_\mathfrak{x}$ liegende Bewertungssystem von $\mathfrak{R}|\mathsf{K}$. Ersichtlich ist $\bigcap_{\mathfrak{b}\in\mathfrak{s}_\mathfrak{x}} \mathfrak{b} = \mathsf{K}[\mathfrak{x}]$, $\bigcap_{B\in\mathfrak{S}_\mathfrak{x}} \mathfrak{B} =$ Ring aller von $\mathsf{K}[\mathfrak{x}]$ ganzabhängigen Elemente $\in\mathfrak{R}$, und $\mathfrak{x}$ ist B-zulässig, genau wenn $B\in\mathfrak{S}_\mathfrak{x}$ ist.

Ferner sei $b_{\mathfrak{x},\infty}$ die durch $b_{\mathfrak{x},\infty}(F(\mathfrak{x})) = -\operatorname{Grad} F$ (F Polynom über K) definierte „Gradbewertung" von $\mathsf{K}(\mathfrak{x})|\mathsf{K}$, $\mathfrak{s}_{|\mathfrak{x}|} = \mathfrak{s}_\mathfrak{x} \cup \{b_{\mathfrak{x},\infty}\}$, $\mathfrak{S}_{|\mathfrak{x}|}$ das über $\mathfrak{s}_{|\mathfrak{x}|}$ liegende Bewertungssystem von $\mathfrak{R}|\mathsf{K}$ ($\mathfrak{S}_{|\mathfrak{x}|} - \mathfrak{S}_\mathfrak{x}$ ist endlich!). Jedem $b\in\mathfrak{s}_{|\mathfrak{x}|}$ wird als *natürlicher Grad (bezüglich* $\mathfrak{x}$) $\mathfrak{f}_{|\mathfrak{x}|}(b)$ der Grad von P (für $b = b_{\mathfrak{x},P}$) bzw. 1 (für $b = b_{\mathfrak{x},\infty}$) zugeordnet; $\mathfrak{F}_{|\mathfrak{x}|}(B) = g_{\mathfrak{R}|\mathsf{K}(\mathfrak{x})} \cdot \mathfrak{f}_{|\mathfrak{x}|}(b)$ heißt der *natürliche Grad (bezüglich* $\mathfrak{x}$) von $B\in\mathfrak{S}_{|\mathfrak{x}|}$. $\mathfrak{s}_{|\mathfrak{x}|}$ bzw. $\mathfrak{S}_{|\mathfrak{x}|}$ heißt das *zu* $\mathfrak{x}$ *gehörige projektive Bewertungssystem* von $\mathsf{K}(\mathfrak{x})|\mathsf{K}$ bzw. $\mathfrak{R}|\mathsf{K}$, weil es invariant gegenüber $P_{\mathfrak{R}|\mathsf{K}}$ ist, d.h. $\mathfrak{s}_{|\Omega\mathfrak{x}|} = \mathfrak{s}_{|\mathfrak{x}|}$, $\mathfrak{S}_{|\Omega\mathfrak{x}|} = \mathfrak{S}_{|\mathfrak{x}|}$; dasselbe gilt auch für die natürlichen Grade: $\mathfrak{f}_{|\Omega\mathfrak{x}|}(b) = \mathfrak{f}_{|\mathfrak{x}|}(b)$, $\mathfrak{F}_{|\Omega\mathfrak{x}|}(B) = \mathfrak{F}_{|\mathfrak{x}|}(B)$ ($b\in\mathfrak{s}_{|\mathfrak{x}|}$, $B\in\mathfrak{S}_{|\mathfrak{x}|}$, $\Omega\in P_{\mathfrak{R}|\mathsf{K}}$).

Bekanntlich gilt[16]):

(6.2) Satz. *$\mathfrak{S}_{|\mathfrak{x}|}$ ist eine unendliche Menge (aber nur für $n=1$ die Menge aller) maxdim. Bewertungen von $\mathfrak{R}|\mathsf{K}$ derart, daß*

a) $\bigcap_{B\in\mathfrak{S}_{|\mathfrak{x}|}} \mathfrak{B} = \mathsf{K}_\mathfrak{R}$ *ist;*

b) $\{B\in\mathfrak{S}_{|\mathfrak{x}|} \mid B(a)\neq 0\}$ *endlich ist für jedes* $a\in\mathfrak{R}$ *(„Endlichkeitsbedingung");*

c) $\sum_{B\in\mathfrak{S}_{|\mathfrak{x}|}} \mathfrak{F}_{|\mathfrak{x}|}(B)\cdot B(a) = 0$ *für jedes* $a\in\mathfrak{R}$ *gilt („Geschlossenheit");*

d) *jedes über einer echten Teilmenge von $\mathfrak{s}_{|\mathfrak{x}|}$ liegende Bewertungssystem unabhängig ist (z.B. $\mathfrak{S}_\mathfrak{x}$).*

Bezeichnet $\mathfrak{s}^{(i)}_{x_i}$ die Menge aller Polynombewertungen von $\mathsf{K}(\mathfrak{x})|\mathsf{K}(\mathfrak{x}^{(i)})$, so ist $\mathfrak{s}_\mathfrak{x} = \bigcup_{i=1\ldots n} \mathfrak{s}^{(i)}_{x_i}$ und damit $\mathfrak{S}_\mathfrak{x} = \bigcup_{i=1\ldots n} \mathfrak{S}^{(i)}_{x_i}$; daher sind die Bewertungssysteme $\mathfrak{S}_{|\mathfrak{x}|}$ und $\bigcup_{i=1\ldots n} \mathfrak{S}^{(i)}_{x_i}$ fast gleich, d.h. sie unterscheiden sich nur um endlich viele Bewertungen. Da $\mathfrak{S}^{(i)}_{|x_i|}$ gleich der Menge aller Bewertungen von $\mathfrak{R}|\mathsf{K}(\mathfrak{x}^{(i)})$ ist, lassen sich die für beliebige a.F.1V. gültigen Aussagen von der Form „für fast alle Bewertungen gilt ..." auf projektive Bewertungssysteme des a.F. nV. $\mathfrak{R}|\mathsf{K}$ übertragen, z.B. Ch. IV § 3, Lemma 3:

(6.3) *Ist $\mathfrak{L}|\mathfrak{R}$ endlich-algebraisch, so ist jede Basis von $\mathfrak{L}|\mathfrak{R}$ B-Ganzheitsbasis von $\mathfrak{L}|\mathfrak{R}$ für fast alle $B\in\mathfrak{S}_{|\mathfrak{x}|}$.*

Eine Abbildung $\mathfrak{A}$ von einem Bewertungssystem $\mathfrak{S}$ von $\mathfrak{R}|\mathsf{K}$ in die Menge der ganzen Zahlen heißt ein $\mathfrak{S}$-*Divisor*; insbesondere heißt $\mathfrak{A}$ *endlich* bzw. *ganz*, wenn $\{B\in\mathfrak{S} \mid \mathfrak{A}B\neq 0\}$ endlich bzw. für alle $B\in\mathfrak{S}$ $\mathfrak{A}B\geq 0$ ist. Ist $\{0\}\subsetneqq M\subseteqq\mathfrak{R}\cup\boldsymbol{d}_{\mathfrak{R}|\mathsf{K}}$, so bezeichnet $\mathfrak{D}_\mathfrak{S}M$ den durch $(\mathfrak{D}_\mathfrak{S}M)B = \min\{B(a) \mid a\in M\}$ ($B\in\mathfrak{S}$) definierten $\mathfrak{S}$-*Divisor von* M. $\mathsf{M}\mathfrak{A} = \{a\in\mathfrak{R} \mid B(a)\geq\mathfrak{A}B \ (B\in\mathfrak{S})\}$ heißt der *zu* $\mathfrak{A}$ *gehörige Modul*. Wir zeigen:

(6.4) Satz. *Für jeden endlichen $\mathfrak{S}_{|\mathfrak{x}|}$-Divisor $\mathfrak{A}$ ist $\mathsf{M}\mathfrak{A}$ ein endlicher K-Modul.*

Beweis. Es sei $\mathfrak{A}_1$ bzw. $\mathfrak{A}_2$ der durch $\mathfrak{A}$ induzierte $(\mathfrak{S}_{|\mathfrak{x}|} - \mathfrak{S}_\mathfrak{x})$- bzw. $\mathfrak{S}_\mathfrak{x}$-Divisor. Dann gibt es $c_1, c_2\in\mathsf{K}[\mathfrak{x}]$ derart, daß $c_1^{-1}\cdot\mathsf{M}\mathfrak{A}_1\subseteqq\mathfrak{J}_{\mathfrak{x},\infty}$ und

[16]) Der in Kr. § 6 nur für separierende $\mathfrak{x}$ geführte Beweis gilt für beliebige Tr.-B. $\mathfrak{x}$. [(2.4) b) gilt ohne Separabilitätsvoraussetzung!]

$c_2 \cdot \mathfrak{M}\mathfrak{A}_2 \subseteq \mathfrak{J}_{\mathfrak{s}_\mathfrak{x}} = \bigcap_{b_\mathfrak{c} \in \mathfrak{s}_\mathfrak{x}} \mathfrak{J}_\mathfrak{c}$ ist. Es sei $\langle v_1, \ldots, v_k \rangle$ eine $b_{\mathfrak{x},\infty}$-Ganzheitsbasis; dann gibt es wegen (2.4) c) und (6.3) ein $c_3 \in \mathsf{K}[\mathfrak{x}]$ derart, daß für jedes $y = \sum_1^k a_i v_i \in \mathfrak{J}_{\mathfrak{s}_\mathfrak{x}}$ $(a_i \in \mathsf{K}(\mathfrak{x}))$ gilt: $c_3 \cdot a_i \in \mathsf{K}[\mathfrak{x}]$ $(1 \leq i \leq k)$. Hieraus folgt, daß sich jedes $z \in \mathfrak{M}\mathfrak{A} = \mathfrak{M}\mathfrak{A}_1 \cap \mathfrak{M}\mathfrak{A}_2$ als Summe $(c_3 \cdot c_2)^{-1} \cdot \sum_1^k D_i(\mathfrak{x}) v_i$ schreiben läßt, wobei Grad $D_i \leq \sum_1^3 j$ Grad C_j $[1 \leq i \leq k,\ C_j(\mathfrak{x}) = c_j]$ ist, also ist $\mathfrak{M}\mathfrak{A}$ in dem durch $\left\{(c_3 \cdot c_2)^{-1} \cdot x_1^{i_1} \cdot \ldots \cdot x_n^{i_n} \cdot v_i \mid i_1 + \cdots + i_n \leq \sum_1^3 j \text{ Grad } C_j,\ 1 \leq i \leq k\right\}$ erzeugten endlichen K-Modul enthalten.

Im folgenden sei $\mathfrak{R} \mid \mathfrak{k}$ endlich-separabel, $[\mathfrak{R} : \mathfrak{k}] = k$, B bzw. b eine maxdim. Bewertung von $\mathfrak{R} \mid \mathsf{K}$ bzw. $\mathfrak{k} \mid \mathsf{K}$. In § 5 haben wir die B-Differente von $\mathfrak{R} \mid \mathfrak{k}$ definiert und für sie den „Exponenten-" und „Schachtelungssatz" [(5.1), (5.2)] bewiesen. Ähnliche Sätze [17]) gelten für die *b-Diskriminante* von $\mathfrak{R} \mid \mathfrak{k}$: $\mathrm{disk}_{\mathfrak{R}\mid\mathfrak{k}} b = \sum_1^r g^{(i)} \cdot \mathrm{Diff}_{\mathfrak{R}\mid\mathfrak{k}} B^{(i)}$ (≥ 0), die mit der üblichen Diskriminante $\Delta_{\mathfrak{R}\mid\mathfrak{k}}(v_1, \ldots, v_k) = |\mathrm{Sp}_{\mathfrak{R}\mid\mathfrak{k}} v_i \cdot v_j|_{1 \ldots k}$ auf folgende Weise zusammenhängt (CH. IV § 8, Lemma 2):

(6.5) $\mathrm{disk}_{\mathfrak{R}\mid\mathfrak{k}} b \leq b\left(\Delta_{\mathfrak{R}\mid\mathfrak{k}}(v_1, \ldots, v_k)\right)$ $(v_1, \ldots, v_k \in \mathfrak{J} = \mathfrak{B}^{(1)} \cap \cdots \cap \mathfrak{B}^{(r)})$; *das Gleichheitszeichen gilt, genau wenn* $\langle v_1, \ldots, v_k \rangle$ *eine b-Ganzheitsbasis von* $\mathfrak{R} \mid \mathfrak{k}$ *ist.*

Für ein Bewertungssystem $\mathfrak{s}$ von $\mathfrak{k} \mid \mathsf{K}$ bzw. $\mathfrak{S}$ von $\mathfrak{R} \mid \mathsf{K}$ bezeichnen wir als $\mathfrak{s}$-*Diskriminante* bzw. $\mathfrak{S}$-*Differente* von $\mathfrak{R} \mid \mathfrak{k}$ den durch $\mathrm{disk}_{\mathfrak{R}\mid\mathfrak{k};\mathfrak{s}} b = \mathrm{disk}_{\mathfrak{R}\mid\mathfrak{k}} b$ $(b \in \mathfrak{s})$ bzw. $\mathrm{Diff}_{\mathfrak{R}\mid\mathfrak{k};\mathfrak{S}} B = \mathrm{Diff}_{\mathfrak{R}\mid\mathfrak{k}} B$ $(B \in \mathfrak{S})$ definierten ganzen $\mathfrak{s}$- bzw. $\mathfrak{S}$-Divisor $\mathrm{disk}_{\mathfrak{R}\mid\mathfrak{k};\mathfrak{s}}$ bzw. $\mathrm{Diff}_{\mathfrak{R}\mid\mathfrak{k};\mathfrak{S}}$.

(6.6) $\mathfrak{s}$ *sei unabhängig und genüge der Endlichkeitsbedingung,* $\mathfrak{S}$ *sei das über* $\mathfrak{s}$ *liegende Bewertungssystem; dann gilt:*

a) $\mathrm{disk}_{\mathfrak{R}\mid\mathfrak{k};\mathfrak{s}} = \mathfrak{d}_\mathfrak{s}\{\Delta_{\mathfrak{R}\mid\mathfrak{k}}(v_1, \ldots, v_k) \mid \langle v_1, \ldots, v_k \rangle \text{ aus } \mathfrak{J}_\mathfrak{s}\}$ $\quad (\mathfrak{J}_\mathfrak{s} = \bigcap_{b_\mathfrak{c} \in \mathfrak{s}} \mathfrak{J}_\mathfrak{c})$.

b) $\mathrm{disk}_{\mathfrak{R}\mid\mathfrak{k};\mathfrak{s}}$ *und* $\mathrm{Diff}_{\mathfrak{R}\mid\mathfrak{k};\mathfrak{S}}$ *sind endlich.*

c) *Fast alle* $B \in \mathfrak{S}$ *sind separabel und unverzweigt bezüglich* $\mathfrak{k}$.

Beweis. a) Wegen (6.5) genügt es zu zeigen, daß zu jedem $b \in \mathfrak{s}$ eine b-Ganzheitsbasis von $\mathfrak{R} \mid \mathfrak{k}$ aus $\mathfrak{J}_\mathfrak{s}$ existiert. — Aus (2.4) a) folgt leicht, daß mit $\mathfrak{s}$ auch $\mathfrak{S}$ der Endlichkeitsbedingung genügt, also ist jede b-Ganzheitsbasis $\langle v_1, \ldots, v_k \rangle$ für fast alle $B_\mathfrak{c} \in \mathfrak{S}$ ein k-Tupel aus $\mathfrak{B}_\mathfrak{c}$, insbesondere für $B^{(1)}, \ldots, B^{(r)}$. Wegen der Unabhängigkeit von $\mathfrak{s}$ ist für eine geeignete Einheit c von $\mathfrak{d}$ $\langle c \cdot v_1, \ldots, c \cdot v_k \rangle$ aus $\mathfrak{J}_\mathfrak{s} = \bigcap_{B_\mathfrak{c} \in \mathfrak{S}} \mathfrak{B}_\mathfrak{c}$ und wegen $\Delta_{\mathfrak{R}\mid\mathfrak{k}}(c \cdot v_1, \ldots, c \cdot v_k) = c^{2k} \cdot \Delta_{\mathfrak{R}\mid k}(v_1, \ldots, v_k)$ nach (6.5) ebenfalls eine b-Ganzheitsbasis.

b) folgt aus a), der Endlichkeitsbedingung und der Ganzheit der Differente.

c) folgt aus b) und (5.1) b).

Ist $\mathfrak{x}$ eine Tr.-B. von $\mathfrak{k} \mid \mathsf{K}$, so erfüllen die über dem Bewertungssystem $\mathfrak{s}_\mathfrak{x}$ (von $\mathsf{K}(\mathfrak{x}) \mid \mathsf{K}$) liegenden Bewertungssysteme $\mathfrak{s}_{\mathfrak{x};\mathfrak{k}}$ (von $\mathfrak{k} \mid \mathsf{K}$) und $\mathfrak{S}_\mathfrak{x}$ (von $\mathfrak{R} \mid \mathsf{K}$) die Voraussetzungen von (6.6) (für $\mathfrak{s}$ bzw. $\mathfrak{S}$). Aus der Endlichkeit von $\mathfrak{S}_{|\mathfrak{x}} - \mathfrak{S}_\mathfrak{x}$ folgt damit der

[17]) Vgl. H. HASSE: Zahlentheorie, S. 322. Berlin: Akademieverlag 1949.

(6.6*) Zusatz. (6.6) b) *und* c) *gelten auch für* $\mathfrak{s} = \mathfrak{s}_{|\mathfrak{x}|;\mathfrak{k}}$, $\mathfrak{S} = \mathfrak{S}_{|\mathfrak{x}|}$ ($\mathfrak{x}$ *Tr.-B. von* $\mathfrak{k}|\mathsf{K}$).

Mit Hilfe von (6.6*) werden wir zeigen, daß im Falle eines vollkommenen Koeffizientenkörpers K zwei (zu separierenden n-Tupeln gehörige) projektive Bewertungssysteme fast gleich sind. Zunächst lassen wir jedoch beliebiges K zu und definieren: es sei ${}^0\mathfrak{s}_{\mathfrak{x}} = \{b_{\mathfrak{x},P} \in \mathfrak{s}_{\mathfrak{x}} | P \text{ separabel}(i) \text{ für ein } i\}$, ${}^0\mathfrak{S}_{\mathfrak{x}}$ das über ${}^0\mathfrak{s}_{\mathfrak{x}}$ liegende Bewertungssystem von $\mathfrak{R}|\mathsf{K}$ ($= \mathfrak{s}_{\mathfrak{x}}$ bzw. $\mathfrak{S}_{\mathfrak{x}}$, falls K vollkommen ist). Aus (2.2) und (5.1) b) folgt unmittelbar:

(6.7) a) $\mathfrak{x}$ *ist B-regulär, genau wenn* $B \in {}^0\mathfrak{S}_{\mathfrak{x}}$ *und B separabel bezüglich* $\mathsf{K}(\mathfrak{x})$ *ist.*

b) $\mathfrak{x}$ *ist B-regulär und B unverzweigt bezüglich* $\mathsf{K}(\mathfrak{x})$, *genau wenn* $B \in {}^0\mathfrak{S}_{\mathfrak{x}}$ *und* $\mathrm{Diff}_{\mathfrak{R}|\mathsf{K}(\mathfrak{x})} B = 0$ *ist.*

Das Bewertungssystem ${}^0\mathfrak{s}_{\mathfrak{x}} \cup \{b_{\mathfrak{x},\infty}\}$ bzw. das darüberliegende von $\mathfrak{R}|\mathsf{K}$ ($= \mathfrak{s}_{|\mathfrak{x}|}$ bzw. $= \mathfrak{S}_{|\mathfrak{x}|}$ für vollkommenes K!) können wir mit ${}^0\mathfrak{s}_{|\mathfrak{x}|}$ bzw. ${}^0\mathfrak{S}_{|\mathfrak{x}|}$ bezeichnen, denn auch

(6.8) ${}^0\mathfrak{s}_{\mathfrak{x}} \cup \{b_{\mathfrak{x},\infty}\}$ *ist invariant bezüglich* $P_{\mathfrak{R}|\mathsf{K}}$.

Beweis. Es sei $\Omega \in P_{\mathfrak{R}|\mathsf{K}}$. Ist $b \in {}^0\mathfrak{s}_{\mathfrak{x}}$, so ist $b = b_{\Omega\mathfrak{x},\infty}$ oder $\in \mathfrak{s}_{\Omega\mathfrak{x}}$. Im letzteren Falle ist nach (6.1) $L_\Omega(\mathfrak{x})$ und damit $D_{\mathfrak{x}}(\Omega\mathfrak{x})$ eine Einheit in $\mathfrak{b}$, also ist mit $\mathfrak{x}$ auch $\Omega\mathfrak{x}$ b-regulär und b-uniformisierend [(6.7) a), (3.4), (3.1)!] und damit $b \in {}^0\mathfrak{s}_{\Omega\mathfrak{x}}$. Wegen $\mathfrak{f}_{|\mathfrak{x}|}(b_{\mathfrak{x},\infty}) = 1$ ist auch $b_{\mathfrak{x},\infty} \in {}^0\mathfrak{s}_{\Omega\mathfrak{x}} \cup \{b_{\Omega\mathfrak{x},\infty}\}$.

(6.9) Satz. *Es seien* $\mathfrak{x}, \mathfrak{y}$ *separierend im a.F.nV.* $\mathfrak{R}|\mathsf{K}$; *dann gilt*:

a) *Ist* $B \in {}^0\mathfrak{S}_{\mathfrak{x}}$, $\mathrm{Diff}_{\mathfrak{R}|\mathsf{K}(\mathfrak{x})} B = 0$, $\mathfrak{y}$ *aus* $\mathfrak{B}$ *und* $B(D_{\mathfrak{x}}\mathfrak{y}) = 0$, *so ist* $B \in {}^0\mathfrak{S}_{\mathfrak{y}}$ *und* $\mathrm{Diff}_{\mathfrak{R}|\mathsf{K}(\mathfrak{y})} B = 0$.

b) ${}^0\mathfrak{S}_{|\mathfrak{x}|}$ *und* ${}^0\mathfrak{S}_{|\mathfrak{y}|}$ *sind fast gleich.*

Beweis. a) Nach (6.7) b) und (3.4) ist B unverzweigt bezüglich $\mathsf{K}(\mathfrak{x})$, $\mathfrak{x}$ B-regulär und B-uniformisierend, und nach (3.1) gilt dasselbe für $\mathfrak{y}$; also folgt a) aus (6.7) b).

b) Wegen (6.6*) und der Endlichkeitsbedingung für ${}^0\mathfrak{S}_{\mathfrak{x}}$ erfüllen fast alle $B \in {}^0\mathfrak{S}_{\mathfrak{x}}$ die Voraussetzungen von a), also sind, da ${}^0\mathfrak{S}_{|\mathfrak{x}|} - {}^0\mathfrak{S}_{\mathfrak{x}}$ endlich ist, fast alle $B \in {}^0\mathfrak{S}_{|\mathfrak{x}|}$ in ${}^0\mathfrak{S}_{|\mathfrak{y}|}$ enthalten.

Als Ergänzung zu § 5 erwähnen wir noch: Ist $B \in (\mathfrak{S}_{|\mathfrak{x}|} - \mathfrak{S}_{\mathfrak{x}}) \cap \mathfrak{S}_{\Omega\mathfrak{x}}$ (für ein $\Omega \in P_{\mathfrak{R}|\mathsf{K}}$), so ist wegen (6.1) c) $B(d\mathfrak{x}) = B(d\Omega_{\mathfrak{x}}) - (n+1) \cdot e_{\mathfrak{x},\infty}$. Da nach (6.6*) und (6.7) a) $\mathfrak{x}$ B-regulär für fast alle $B \in {}^0\mathfrak{S}_{|\mathfrak{x}|}$ ist, stimmen wegen (5.7) und (5.9) die Divisoren $\mathfrak{D}_{\mathfrak{S}_{|\mathfrak{x}|}}\{d\mathfrak{x}\}$ und $\mathrm{Diff}_{\mathfrak{R}|\mathsf{K}(\mathfrak{x});\,\mathfrak{S}_{|\mathfrak{x}|}} \cdot (\mathrm{Con}_{\mathfrak{R}|\mathsf{K}(\mathfrak{x})} \mathfrak{p}_{\mathfrak{x},\infty})^{-(n+1)}$ [18]) für fast alle $B \in {}^0\mathfrak{S}_{|\mathfrak{x}|}$ ($= \mathfrak{S}_{|\mathfrak{x}|}$, falls K vollkommen) überein; im Falle $Ch(\mathsf{K}) = 0$ sind sie sogar gleich.

[18]) $\mathfrak{p}_{\mathfrak{x},\infty}$ bezeichnet den zu $b_{\mathfrak{x},\infty}$ gehörigen Primdivisor von $\mathfrak{s}_{|\mathfrak{x}|}$, $(\mathrm{Con}_{\mathfrak{R}|\mathfrak{k}}\mathfrak{a})\, B = e_{\mathfrak{R}|\mathfrak{k}} \cdot \mathfrak{a}b$ (vgl. Ch. IV § 7). Multiplikation der Divisoren wie üblich.

Mathematisches Institut der Universität Bonn

(Eingegangen am 8. Dezember 1956)

SCHAUFFLER, R.
Math. Zeitschr. Bd. 67, S. 428—435 (1957)

Die Assoziativität im Ganzen, besonders bei Quasigruppen

Von

RUDOLF SCHAUFFLER

1. M sei die nicht leere endliche Menge $\{1, 2, \ldots, n\}$. In M sei eine geordnete Menge Ω von zweistelligen Operationen in Gestalt von Verknüpfungstafeln gegeben. Eine Tafel $T \in \Omega$ ordnet jedem geordneten Paar $x, y \in M$ ein Bildelement xTy aus M zu[1]). Wir denken uns die Tafeln in der Normalform gegeben, die aus einer quadratischen Matrix für die Variable xTy mit einer Außenspalte (links) für x und einer Außenzeile (oben) für y besteht, wobei die Außenreihen die Zahlen in der natürlichen Reihenfolge enthalten (siehe z. B. Bild 1).

Wir bezeichnen die Tafel $\overline{T}$, die aus T durch Spiegelung an der Hauptdiagonale hervorgeht (so daß $xTy = y\overline{T}x$ für $x, y \in M$), als die zu T *transponierte* Tafel; diese Beziehung zwischen T und $\overline{T}$ ist symmetrisch. Wir nennen Ω *im Ganzen kommutativ*, wenn es zu jeder Tafel $T \in \Omega$ eine transponierte $\overline{T} \in \Omega$ gibt.

Wir nennen zwei geordnete Tafelpaare $T_1 T_2$ und $T_1' T_2'$ *links* bzw. *rechts konjugiert*, wenn $(xT_1y)T_2z = (xT_1'y)T_2'z$ bzw. $xT_1(yT_2z) = xT_1'(yT_2'z)$ für $x, y, z \in M$. Beide Beziehungen sind symmetrisch und transitiv, führen also zu einer Einteilung der Tafelpaare in Klassen, die wir kurz *Linksklassen* und *Rechtsklassen* nennen wollen.

Wenn es zu $T_1 T_2$ ein Paar $\tilde{T}_1 \tilde{T}_2$ gibt, so daß $(xT_1y)T_2z = (z\tilde{T}_1y)\tilde{T}_2x$ für $x, y, z \in M$, so heiße $\tilde{T}_1 \tilde{T}_2$ *links invers* zu $T_1 T_2$; analog wird „*rechts invers*" definiert. Beide Beziehungen sind symmetrisch. Wir nennen Ω *im Ganzen invertierbar*, wenn es in Ω zu jedem Tafelpaar mindestens ein links und ein rechts inverses gibt.

Wenn $(xT_1y)T_2z = xT_1^*(yT_2^*z)$ für $x, y, z \in M$, so nennen wir $T_1^* T_2^*$ *rechts assoziiert* zu $T_1 T_2$, $T_1 T_2$ *links assoziiert* zu $T_1^* T_2^*$. Dementsprechend reden wir auch von rechts oder links assoziierten Tafelpaarklassen. Ist ein Tafelpaar zu einem anderen (eventuell mit ihm identischen) rechts und links assoziiert, so nennen wir die beiden Paare zueinander *assoziiert*; diese Beziehung ist symmetrisch. Ein zu sich selbst assoziiertes Tafelpaar heiße *assoziativ*. Wir nennen Ω *im Ganzen assoziativ*, wenn es in Ω zu jedem Tafelpaar $T_1 T_2$ mindestens ein rechts assoziiertes $T_1^* T_2^*$ und ein links assoziiertes ${}^*T_1 {}^*T_2$ gibt.

[1]) Ω ist also eine Algebra vom Typ $\langle 2, 2, \ldots, 2 \rangle$ im Sinne von H. HERMES, Verbandstheorie (1955), S. 153.

Wenn Ω i. G. kommutativ und invertierbar ist, so gilt:

$$(x\,T_1 y)\,T_2 z = z\,\overline{T}_2\,(y\,\overline{T}_1\,x) = x\,\widetilde{T}_2\,(y\,\widetilde{T}_1\,z),$$

$$x\,T_1(y\,T_2\,z) = (z\,\overline{T}_2\,y)\,\overline{T}_1\,x = (x\,\widetilde{T}_2\,y)\,\widetilde{T}_1\,z \qquad \text{für } x, y, z \in M,$$

d. h. Ω ist i. G. assoziativ. Wenn Ω i. G. kommutativ und assoziativ ist, so gilt:

$$(x\,T_1 y)\,T_2 z = x\,T_1^*(y\,T_2^*\,z) = (z\,\overline{T_2^*}\,y)\,\overline{T_1^*}\,x,$$

$$x\,T_1(y\,T_2\,z) = (x\,{}^*T_1 y)\,{}^*T_2 z = z\,{}^*\overline{T_2}\,(y\,{}^*\overline{T_1}\,x),$$

d. h. Ω ist i. G. invertierbar. Also gilt:

Satz 1. *Ist Ω i. G. kommutativ, so ist Ω genau dann i. G. assoziativ, wenn es i. G. invertierbar ist.*

Bei festem $T_1\,T_2$ verwenden wir die Bezeichnungen $s_{xyz} = (x\,T_1 y)\,T_2 z$, $s^*_{xyz} = x\,T_1(y\,T_2 z)$. Die lexikographisch geordnete Elementenfolge $s_{111}, s_{112}, \ldots, s_{nnn}$ heiße das *linke Wertsystem* zu $T_1\,T_2$, die Folge $s^*_{111}, \ldots, s^*_{nnn}$ das *rechte Wertsystem* zu $T_1\,T_2$. Ω ist genau dann i. G. assoziativ, wenn die Menge der rechten Wertsysteme gleich der Menge der linken Wertsysteme ist.

Beim Rechnen mit einer einzelnen assoziativen Verknüpfungstafel kann man bekanntlich jeden Klammerausdruck auf eine Normalform bringen, an deren Anfang alle „Klammer auf"-Zeichen vereinigt sind. Dasselbe gilt nun auch für eine i. G. assoziative Operationenmenge Ω. So erhält man z. B. aus

$$x_4\,T_1\big((x_1\,T_2\,x_3)\,T_3\,(x_2\,T_4\,x_5)\big) \tag{1}$$

nach drei Umformungsschritten als erste Normalform:

$$\big(((x_4\,U_1\,x_1)\,U_2\,x_3)\,U_3\,x_2\big)\,U_4\,x_5 \tag{2}$$

mit geeigneten $U_i \in \Omega$. Ist Ω auch noch i. G. kommutativ, so kann man beim Umformen auch vom kommutativen Gesetz und wegen Satz 2 von der Invertierbarkeit Gebrauch machen und zu einer zweiten Normalform gelangen, in der die x_i in der normalen Reihenfolge der Indizes stehen. Man erhält so aus (1) die zweite Normalform:

$$\big(((x_1\,V_1\,x_2)\,V_2\,x_3)\,V_3\,x_4\big)\,V_4\,x_5 \tag{3}$$

mit geeigneten $V_i \in \Omega$.

Auch Klammerausdrücke mit einer einzigen, nicht assoziativen Verknüpfungstafel lassen sich auf diese Art in die erste (und zweite) Normalform bringen, wenn man in Kauf nimmt, daß in der Normalform mehrere verschiedene Verknüpfungstafeln auftreten; Voraussetzung ist dabei, daß man ein i. G. assoziatives (und kommutatives) Ω kennt, das die gegebene Tafel enthält.

2. Wenn die quadratische Matrix der Verknüpfungstafel ein lateinisches Quadrat ist, d.h. in jeder Zeile und Spalte jedes Element genau einmal vorkommt, so heißt M im Verein mit der Verknüpfung eine *Quasigruppe n-ter Ordnung*. Eine Verknüpfung ist auch dadurch als Quasigruppe charakterisiert, daß jede der Gleichungen $a\,T\,x = b$ und $y\,T\,a = b$ genau eine Lösung in M besitzt, wenn $a, b \in M$. Die Menge der Quasigruppen mit den Elementen $1, 2, \ldots, n$ werde mit Ω_n bezeichnet. Man erkennt leicht, daß Ω_n i.G. kommutativ ist.

Wie steht es aber bei Quasigruppen mit der Assoziativität i.G.? Diese Frage ist auch von einer gewissen praktischen Bedeutung: Will man für einen Telegraphencode einen Vorrat von *Codewörtern* herstellen, so empfiehlt es sich, die Codewörter, die beispielsweise Buchstaben-Quintupel sind, so zu gestalten, daß sie *bidifferent* werden, d.h. sich je zwei von ihnen an mindestens zwei der fünf Stellen unterscheiden. In diesem Fall und ganz allgemein, wenn man eine Menge von Gegenständen mit bidifferenten Nummern versehen will, kann man den Zweck erreichen, indem man einen Quasigruppen-Klammerausdruck — z.B. von der Form (1) — einer Konstanten gleichsetzt. Als Elemente kann man Buchstaben oder Ziffern wählen. Die Gesamtheit der Lösungsquintupel dieser Gleichung ist dann der Codewortschatz oder das Nummernsystem[2]). Falls nun die Gesamtheit der Quasigruppen der betreffenden Ordnung i.G. assoziativ ist, erhält man, da Ω_n i.G. kommutativ ist, schon bei Beschränkung auf die Klammerausdrücke der zweiten Normalform alle Codewortschätze oder Nummernsysteme dieser Art. — Leider erweist sich diese Möglichkeit als sehr beschränkt; denn wir werden zeigen:

Satz 2. *Ω_n ist dann und nur dann i.G. assoziativ, wenn $n \leq 3$ ist.*

Beweis. Der Fall $n = 1$ ist trivial.

Im Falle $n = 2$ gibt es die beiden Tafeln

$$(D_1)\quad \begin{array}{c|cc} & 1 & 2 \\ \hline 1 & 1 & 2 \\ 2 & 2 & 1 \end{array} \quad \text{und} \quad (D_2)\quad \begin{array}{c|cc} & 1 & 2 \\ \hline 1 & 2 & 1 \\ 2 & 1 & 2 \end{array}$$

und die beiden Tafelpaarklassen D_1D_1, D_2D_2 und D_1D_2, D_2D_1, deren jede zugleich Rechts- und Linksklasse und zu sich selbst assoziiert ist; die erste ergibt als linkes und rechtes Wertsystem: 12212112, die zweite 21121221.

Im Falle $n = 3$ haben wir es mit zwei Typen von Tafeln zu tun: Typ 1, vertreten durch Tafel A_1, und Typ 2, vertreten durch Tafel A_2 von Bild 1.

$$(A_1)\quad \begin{array}{c|ccc} & 1 & 2 & 3 \\ \hline 1 & 1 & 2 & 3 \\ 2 & 2 & 3 & 1 \\ 3 & 3 & 1 & 2 \end{array} \qquad (A_2)\quad \begin{array}{c|ccc} & 1 & 2 & 3 \\ \hline 1 & 1 & 2 & 3 \\ 2 & 3 & 1 & 2 \\ 3 & 2 & 3 & 1 \end{array}$$

Bild

[2]) SCHAUFFLER, R.: Über die Bildung von Codewörtern. Arch. elektr. Übertragung **10**, 303—314 (1956).

Man erhält alle Tafeln von Typ 1 (Typ 2), indem man auf die Ziffern der quadratischen Matrix von A_1 (A_2) die 6 Permutationen der symmetrischen Gruppe $\mathfrak{S}_3$ anwendet, aber die Außenreihen unverändert läßt. Typ 1 ist durch $1\,T\,2 = 2\,T\,1$, Typ 2 durch $1\,T\,2 \neq 2\,T\,1$ charakterisiert.

Es wird sich später als zweckmäßig erweisen, wenn wir für die 12 Tafeln folgende dreistellige Numerierung einführen: Wir numerieren die Permutationen von $\mathfrak{S}_3$ durch Zahlenpaare (s. Bild 2) und numerieren jede Tafel vom Typ 1 bzw. 2, indem wir an die Ziffer 1 bzw. 2 die Nummer der betreffenden Permutation anhängen. Wir erhalten so 12 Tafelnummern ijk ($i, j = 1, 2$; $k = 1, 2, 3$). A_1 und A_2 erhalten so die Nummern 111 bzw. 211. Nur 111, 112, 113 sind Gruppentafeln. Die Paare 211, 221; 212, 222 und 213, 223 sind Paare transponierter Tafeln; die übrigen 6 Tafeln (es sind die Tafeln $1jk$) sind zu sich selbst transponiert.

ν	1	2	3	Nr.
	1	2	3	11
	2	3	1	12
	3	1	2	13
$p(\nu)$	1	2	3	21
	2	1	3	22
	3	2	1	23

Bild 2

Man gelangt nun wie bei $n = 2$ am raschesten zum Nachweis der Assoziativität i. G., wenn man die Wertsysteme effektiv berechnet. Dabei genügt es aber, von jedem Wertsystem nur so viele von den 27 Werten zu berechnen, daß man bei je zwei Wertsystemen entscheiden kann, ob sie gleich oder verschieden sind.

Wir wollen zeigen: Bei $n = 3$ ist jedes linke Wertsystem durch das Wertequadrupel $s_1 = s_{111}$, $s_2 = s_{112}$, $s_3 = s_{121}$, $s_4 = s_{211}$ und jedes rechte durch $s_1^* = s_{111}^*$, ..., $s_4^* = s_{211}^*$ eindeutig bestimmt. Zum Beweis benützen wir den Umstand, daß ein lateinisches Quadrat 3. Ordnung durch drei seiner Elemente bestimmt ist, von welchen zwei in einer Zeile und zwei in einer Spalte stehen. Es ist leicht zu sehen, daß die fünf Matrizen $S_1 = (s_{1yz})$, $S_2 = (s_{x1z})$, $S_3 = (s_{xy1})$, $S_4 = (s_{2yz})$, $S_5 = (s_{3yz})$ $(x, y, z = 1, 2, 3)$ lateinische Quadrate sind. In S_1 werden durch s_{111}, s_{112} und s_{121} 6 weitere Werte bestimmt; in S_2 sind dann die 4 Werte s_{111}, s_{112}, s_{113}, s_{211} bekannt, wodurch 5 neue Werte bestimmt werden. Aus S_3, S_4 und S_5 ergeben sich noch je 4 Werte, so daß nun alle 27 Werte bestimmt sind. Derselbe Beweis gilt auch für die rechten Wertsysteme.

Ich habe nun für die 12^2 Tafelpaare die Wertequadrupel berechnet und festgestellt: Es gibt 24 verschiedene Quadrupel $s_1 s_2 s_3 s_4$; zu jedem Quadrupel gehört eine Linksklasse und eine Rechtsklasse von je 6 Tafelpaaren. Damit ist erwiesen, daß Ω_3 i. G. assoziativ ist.

3. Ich will nun den Beweis von Satz 2 unterbrechen, um einige der Gesetzmäßigkeiten der Algebra Ω_3 mitzuteilen, die sich bei dieser Gelegenheit erkennen ließen. Dabei erweist es sich als zweckmäßig, auch für die Wertequadrupel gewisse dreistellige Nummern einzuführen.

Die 24 Quadrupel zerfallen in vier Typen:

Typ Nummer	1,	2,	3,	4,
vertreten durch	1222,	1233,	1223,	1232.

Man erhält die 6 Quadrupel jeden Typs, indem man auf die Ziffern des hier angegebenen Vertreters die 6 Permutationen von $\mathfrak{S}_3$ anwendet. Wir numerieren die 6 Quadrupel jeden Typs, indem wir an die Nummer des Typs die Nummer der betreffenden Permutationen aus Bild 2 anhängen. So erhalten wir für jedes Quadrupel (und damit auch für jedes Wertsystem) eine Nummer der Form uvw $(u=1, 2, 3, 4;\ v=1, 2;\ w=1, 2, 3)$.

Wir vermerken für später, daß zu jedem Quadrupel $uvw = s_1 s_2 s_3 s_4$ auch das Quadrupel $\tilde{u}\tilde{v}\tilde{w} = s_1 s_4 s_3 s_2$ vorhanden ist, was auch aus Satz 1 folgt. Man erhält

$$(4)\quad \begin{cases} \text{für } u=1 \text{ und } u=4, \text{ d.h. } s_2 = s_4\colon\ \tilde{u}\tilde{v}\tilde{w} = uvw; \\ \text{für } u=2 \text{ und } u=3, \text{ d.h. } s_2 \neq s_4\colon\ \tilde{u}=5-u,\ \tilde{v}\equiv v+1 \pmod 2,\ \tilde{w}=w. \end{cases}$$

$T_2 = i_2 j_2 k_2$

$T_1 = i_1 j_1 k_1$	111	112	113	121	122	123	211	212	213	221	222	223	uvw
111	111	2	3	121	2	3	211	2	3	221	2	3	
112	2	3	1	3	1	2	3	1	2	2	3	1	
113	3	1	2	2	3	1	2	3	1	3	1	2	
121	211	2	3	221	2	3	111	2	3	121	2	3	
122	2	3	1	3	1	2	3	1	2	2	3	1	
123	3	1	2	2	3	1	2	3	1	3	1	2	
211	311	2	3	321	2	3	411	2	3	421	2	3	
212	2	3	1	3	1	2	3	1	2	2	3	1	
213	3	1	2	2	3	1	2	3	1	3	1	2	
221	411	2	3	421	2	3	311	2	3	321	2	3	
222	2	3	1	3	1	2	3	1	2	2	3	1	
223	3	1	2	2	3	1	2	3	1	3	1	2	

Bild 3

Bild 3 zeigt die errechnete Tabelle, die jedem Tafelpaar sein linkes Wertsystem zuordnet. Sie ergibt z.B. zu $T_1\,T_2 = 211\,121$ bzw. $212\,123$ das linke Wertsystem 321 bzw. 322.

Bild 4

$i_1 j_1$ \ i_2	1	2
11	1	2
12	2	1
21	3	4
22	4	3

u

j_2	v
1	1
2	2

für $i_2 + j_2 \equiv 0$

k_1 \ k_2	1	2	3
1	1	2	3
2	2	3	1
3	3	1	2

w

für $i_2 + j_2 \equiv 1 \pmod 2$

k_1 \ k_2	1	2	3
1	1	2	3
2	3	1	2
3	2	3	1

w

Bild 3 läßt sich auf Bild 4 reduzieren und alsdann durch folgende Formeln ersetzen:

$$(5)\quad \begin{cases} u = 2i_1 - \tfrac{1}{2}\left(1 + (-1)^{i_1+i_2}\right), \quad v = j_2, \\ w = k_1 A_\nu k_2 \text{ mit } \nu \equiv i_2 + j_2 - 1 \pmod 2 \quad \text{(wegen } A_\nu \text{ s. Bild 1)} \\ \text{oder } w \equiv (k_1 - 1)(-1)^{i_2+j_2} + k_2 \pmod 3. \end{cases}$$

Aus (5) oder Bild 3 läßt sich Folgendes herleiten: Zwei Tafelpaare $i_1 j_1 k_1$ $i_2 j_2 k_2$ und $i_1'' j_1'' k_1'$ $i_2'' j_2'' k_2'$ sind genau dann links konjugiert, wenn

$$(6)\qquad \begin{cases} i_1'' = i_1, \quad i_2' - i_2 \equiv j_1'' - j_1 \pmod 2, \quad j_2'' = j_2, \\ k_2' - k_2 \equiv (k_1 - 1)(-1)^{i_2+j_2} - (k_1' - 1)(-1)^{i_2''+j_2''} \pmod 3. \end{cases}$$

Um nach (6) zu dem Tafelpaar $i_1 j_1 k_1$ $i_2 j_2 k_2$ ein links konjugiertes zu erhalten, wählt man j_1'' und k_1' beliebig und erhält dann aus (6) die Elemente i_1'', i_2'', j_2'', k_2'; so ergibt sich eine Linksklasse von 2.3 Tafelpaaren (Beispiel: Bild 5a und c).

Bild 5

1233				1332			
111	211	121	221	211	121	221	111
2	2	2	2	2	2	2	2
3	3	3	3	3	3	3	3
121	111	211	211	221	221	111	121
2	3	3	2	2	3	3	2
3	2	2	3	3	2	2	3
a		b		c		d	

Nun werde gezeigt, wie man zu einer Linksklasse die rechts assoziierte Klasse ermittelt. Da Ω_3 i. G. kommutativ, assoziativ und invertierbar ist, folgt für $T_1, T_2 \in \Omega_3$:

$$(x\,T_1\,y)\,T_2\,z = x\,T_1^*\,(y\,T_2^*\,z),$$
$$(x\,T_1\,y)\,T_2\,z = (z\,\widetilde{T}_1\,y)\,\widetilde{T}_2\,x,$$

also

$$x\,T_1^*\,(y\,T_2^*\,z) = (z\,\widetilde{T}_1\,y)\,\widetilde{T}_2\,x = x\,\overline{\widetilde{T}}_2\,(y\,\overline{\widetilde{T}}_1\,z) \qquad \text{für } x, y, z \in M;$$

d. h. man erhält zu der zum Quadrupel $s_1 s_2 s_3 s_4$ gehörigen Linksklasse die rechts assoziierte Rechtsklasse, indem man mit Hilfe von (4) und Bild 3 die Tafelpaare der Linksklasse zu $s_1 s_4 s_3 s_2$ ermittelt, die beiden Tafeln jedes Paares vertauscht und jede Tafel durch die transponierte ersetzt.

Beispiel. Bild 5 zeigt die Linksklasse a bzw. c und die Rechtsklasse b bzw. d zu Quadrupel $\underline{12}\underline{33}$ (Nummer 211) bzw. $\underline{13}\underline{32}$ (321).

In jeder Linksklasse gibt es ein Tafelpaar $T_{01}\,T_{02} = i_{01}\,1\,1\;i_{02}\,j_{02}\,k_{02}$; die Linksklasse ist durch das Paar $i_{01}\,T_{02}$ oder das Quadrupel $i_{01}, 1\,T_{02}1, 1\,T_{02}2, 2\,T_{02}1$ eindeutig bestimmt und es besteht die merkwürdige Beziehung:

$$(7)\qquad \begin{cases} s_1 = 1\,T_{02}1, \; s_2 = 1\,T_{02}2, \; s_3 = 2\,T_{02}1, \\ s_4 = \begin{cases} s_3 & \text{für } i_{01} = 1, \\ 6 - s_1 - s_3 & \text{für } i_{01} = 2. \end{cases} \end{cases}$$

Beispiel. Die Tafelpaare 111 211 bzw. 211 121 in Bild 5a bzw. c (Tafel 211 ist A_2 von Bild 1).

Wählt man für die Menge der Gruppen n-ter Ordnung mit Elementen aus M die Bezeichnung Γ_n, so zeigt es sich, daß $\Gamma_3 = \{111, 112, 113\}$ i. G. assoziativ

ist; es ist sogar jedes Tafelpaar aus Γ_3 assoziativ. Eine andere i.G. assoziative Teilmenge von Ω_3 ist z.B. $\{111, 121, 211, 221\}$.

4. Nach diesem Exkurs kehren wir zum Beweis von Satz 2 zurück. Wir betrachten die Gleichung

$$(x_1 T_1 x_3)\, T_2\, x_2 = x_4 \tag{8}$$

und weisen nach, daß bei $n \geqq 4$ und passend gewähltem $T_1 T_2$ die Lösungsgesamtheit dieser Gleichung eine Struktureigenschaft hat, die bei einem i.G. assoziativen Ω_n nicht möglich ist.

Wenn Ω_n i.G. assoziativ wäre, so ließe sich (8) wie folgt umformen:

$$x_1 U_1 (x_3 U_2 x_2) = x_4, \quad x_1 U_1 (x_2 \overline{U}_2 x_3) = x_4, \quad (x_1 V_1 x_2) V_2 x_3 = x_4,$$
$$x_1 V_1 x_2 = f(x_3, x_4),$$

wobei $f(x_3 x_4)$ für $x_3, x_4 \in M$ n verschiedene Werte annehmen könnte. Da die letzte Gleichung bei festem x_1, $f(x_3, x_4)$ und bei festem x_2, $f(x_3, x_4)$ eine eindeutige Lösung x_2 bzw. x_1 hat, würde man bei festem $f(x_3, x_4)$ als Lösungspaare $x_1 x_2$ die n Paare einer Permutation $x_1 \to x_2$ erhalten, und bei allen möglichen Werten des Paares $x_3 x_4$ genau n verschiedene Permutationen $x_1 \to x_2$. Wir zeigen nun aber, daß es bei $n \geqq 4$ stets Tafelpaare von der Art gibt, daß bei variierendem $x_3 x_4$ mehr als n verschiedene Permutationen $x_1 \to x_2$ auftreten. Man erhält die beiden Matrizen eines solchen Tafelpaares $T_1 T_2$, indem man in der Matrix (a_{ik}), wo $a_{ik} \in M$, $a_{ik} \equiv i + k - 1 \pmod n$, die beiden ersten Spalten bzw. Zeilen vertauscht (Bild 6). Nun läßt sich an Hand von

	1	2	3	4	5
1	**2**	**1**	3	4	5
2	**3**	**2**	4	5	1
3	**4**	**3**	5	1	2
4	**5**	**4**	1	2	3
5	**1**	**5**	2	3	4

T_1

	1	2	3	4	5
1	**2**	**3**	**4**	**5**	**1**
2	**1**	**2**	**3**	**4**	**5**
3	3	4	5	1	2
4	4	5	1	2	3
5	5	1	2	3	4

T_2

Bild 6

x_3	x_4	x_1	x_2	x_1	x_2
		1	$p(1)$	2	$p(2)$
1	1	1	1	2	4
1	2	1	2	2	5
1	3	1	3	2	1
1	4	1	4	2	2
1	5	1	5	2	3
2	1	1	5	2	1

Bild 7

(8) und Bild 6 eine Tabelle aufstellen, die zu den $n+1$ Paaren $x_3 x_4 = 11$, $12, \ldots, 1n$, 21 die zugehörigen Permutationen $x_1 \to x_2$ oder $x_1 \to p(x_1)$ angibt. Dabei genügt es, von jeder Permutation die beiden Anfangspaare 1, $p(1)$ und 2, $p(2)$ zu notieren (Bild 7). Man erhält für $x_3 x_4 = 11, \ldots, 1n$ n verschiedene Permutationen $x_1 \to x_2$, deren i-te mit dem Paar $x_1 x_2 = 1i$ beginnt, worauf bei der n-ten das Paar $x_1 x_2 = 2(n-2)$ folgt. Für $x_3 x_4 = 21$ dagegen ergibt sich eine Permutation $x_1 \to x_3$, die mit $x_1 x_2 = 1n$ und 21 beginnt, also bei $n > 3$ von den obigen n Permutationen verschieden ist. Damit ist Satz 2 bewiesen.

5. Da Ω_n für $n \geq 4$ nicht i. G. assoziativ ist, ergibt sich die Aufgabe, die i. G. assoziativen Teilmengen von Ω_n zu ermitteln. Insbesondere erhebt sich die Frage, wie es bei Γ_n, der Menge aller Gruppen aus Ω_n, mit der Assoziativität i. G. steht[3]). Für $n=4$ habe ich einstweilen Folgendes festgestellt: Zu dem Tafelpaar von Bild 8 gibt es kein links assoziiertes; also ist Γ_4 nicht i. G. assoziativ. Γ_4 besteht bekanntlich aus 4 nicht-zyklischen (zu Bild 8 a isomorphen) und 12 zyklischen (zu Bild 8 b isomorphen) Gruppen. Die Menge der 4 nicht-zyklischen Gruppen ist i. G. assoziativ, weil hier jedes einzelne Tafelpaar assoziativ ist.

	1	2	3	4
1	1	2	3	4
2	2	1	4	3
3	3	4	1	2
4	4	3	2	1

a

	1	2	3	4
1	2	3	4	1
2	3	4	1	2
3	4	1	2	3
4	1	2	3	4

b

Bild 8

Zum Schluß sei noch ein sehr bekanntes Ω erwähnt, das nicht aus Quasigruppen besteht: die Boolesche Algebra mit zwei Elementen:

	1	2
1	1	2
2	2	2

	1	2
1	1	1
2	1	2

.

Man findet hier leicht, daß es 6 Wertsysteme gibt, von welchen zwei nur rechte und zwei nur linke Wertsysteme sind. Dieses Ω ist also nicht i. G. assoziativ.

[3]) Die Gruppen sind bekanntlich genau die assoziativen Quasigruppen.

Urach, Münsinger Str. 4

(Eingegangen am 15. Dezember 1956)

SALZMANN, H.
Math. Zeitschr. Bd. 67, S. 436—466 (1957)

Topologische projektive Ebenen

Von
HELMUT SALZMANN

Einleitung

Unter einer topologischen projektiven Ebene verstehen wir eine projektive Ebene, deren Punkt- und Geradenmenge topologische Räume sind derart, daß Verbinden und Schneiden stetige Operationen sind. Solche Ebenen haben eine Reihe überraschender topologischer Eigenschaften; z.B. sind die zugehörigen topologischen Räume regulär. Darüber hinaus können wir in Beantwortung einer von FREUDENTHAL [**1956**] gestellten Frage zeigen, daß eine lokalkompakte projektive Ebene das zweite Abzählbarkeitsaxiom erfüllt. Die Geraden wie auch die punktierten Geraden einer zusammenhängenden lokalkompakten Ebene erweisen sich als kurvenweise zusammenhängend und lokal kurvenweise zusammenhängend.

Soweit gelten die Sätze in einer beliebigen topologischen projektiven Ebene, ohne daß irgendwelche Schließungssätze vorausgesetzt werden. Unter der Voraussetzung gewisser einfacher Schließungssätze läßt sich für die Topologie der Ebene noch mehr folgern: Gilt der desarguessche Satz für eine feste Achse und ein festes nicht auf ihr gelegenes Zentrum, so ist eine lokalkompakte projektive Ebene sogar kompakt. Dasselbe gilt übrigens auch, wenn die Topologie durch eine Anordnung der Ebene hervorgerufen wird. (SKORNJAKOV hatte früher [**1954**] bewiesen, daß man in einer zusammenhängenden projektiven Ebene von der lokalen Kompaktheit auf die Kompaktheit der ganzen Ebene schließen kann.) Ist für ein Parallelgewebe in der Ebene die Thomson-Bedingung erfüllt (das bedeutet, daß die zu dem Gewebe gehörige Loop eine kommutative Gruppe ist), so sind die Geraden der Ebene homöomorph zu einer Sphäre; insbesondere hat die Ebene dann eine endliche topologische Dimension. Nebenbei ergibt sich bei dem Beweis, daß in einer zusammenhängenden kompakten Ebene nicht jedes vollständige Viereck kollineare Diagonalpunkte haben kann.

Der letzte Teil der Arbeit widmet sich den projektiven Ebenen, die zur gewöhnlichen reellen Ebene homöomorph sind. Wir zeigen zunächst, daß in einer solchen Ebene auch die Geraden homöomorph zur reellen Geraden sind. Man wird nicht erwarten können, daß die topologische Voraussetzung der Homöomorphie zur reellen Ebene allein die allgemeine Gültigkeit irgendeines Schließungssatzes zur Folge hat. Tatsächlich erweist es sich, wie in einer späteren Arbeit gezeigt werden soll, daß in der [**1902**] von MOULTON konstruierten nichtdesarguesschen Ebene mit der gewöhnlichen reellen Topologie kein Schließungssatz allgemein gültig ist. Jedoch bestehen in einer zur reellen Ebene homöomorphen projektiven Ebene viel engere Beziehungen

zwischen Schließungssätzen, als ohne die topologische Voraussetzung. So ist etwa eine solche Ebene, in welcher der desarguessche Satz für eine feste Achse und jedes auf ihr gelegene Zentrum gilt (Translationsebene), sogar gleich der reellen Ebene. Erstaunlich ist neben diesem Resultat die Existenz einer projektiven Ebene mit der Topologie der reellen Ebene, in welcher der desarguessche Satz genau für die drei Ecken eines festen Dreiecks als Zentren und die Gegenseiten als zugehörigen Achsen gilt. Als Koordinatenbereich hat diese Ebene einen (echten, planaren, kommutativen) Neokörper mit kürzbarer Addition, womit die von PICKERT [**1956**] aufgeworfene Frage nach der Existenz solcher Neokörper beantwortet ist.

In einem angehängten Paragraphen wird die Frage, ob ein topologischer Ternärkörper Koordinatenbereich einer topologischen projektiven Ebene ist, für Alternativkörper bejahend beantwortet.

Den Herren R. BAER und G. PICKERT fühle ich mich für ihre vielen und wertvollen Anregungen zu großem Dank verpflichtet.

§ 1. Definitionen

Eine projektive Ebene besteht aus zwei Mengen von Dingen, die Punkte und Geraden genannt werden, und einer als Enthaltensein bezeichneten Relation zwischen Punkten und Geraden, die den folgenden drei Bedingungen genügt:

(I) Zu zwei verschiedenen Punkten p, q gibt es genau eine Gerade $G = p \cup q$, in der die beiden Punkte enthalten sind, $p \in G$ und $q \in G$. G heißt die Verbindungsgerade von p und q.

(II) Dual dazu gibt es zu zwei verschiedenen Geraden G, H genau einen „Schnittpunkt" $p = G \cap H$, der in beiden Geraden enthalten ist, $p \in G$ und $p \in H$.

(III) Es gibt vier verschiedene Punkte, von denen keine drei in ein und derselben Geraden enthalten sind.

Statt „p ist in der Geraden G enthalten" sagt man auch: p liegt auf G, G geht durch p, der Punkt p gehört der Geraden G an, G enthält p. Man kann eine Gerade G mit der Menge der in ihr enthaltenen Punkte identifizieren. Bei dieser Auffassung ist also jede Gerade eine Teilmenge der Gesamtheit P aller Punkte. $\mathfrak{G}$ sei die Menge aller Geraden[1]). Eine Menge P wird zu einem topologischen Raum, wenn man gewisse Teilmengen von P als offene Mengen auszeichnet, derart, daß jede Vereinigung von offenen Mengen wieder offen ist (auch die leere Menge $\emptyset$ als „leere Vereinigung") und ebenso jeder Durchschnitt von endlich vielen offenen Mengen (darunter auch der ganze Raum P als „leerer Durchschnitt"). Die Komplemente offener Mengen heißen abgeschlossen. Umgebung eines Punktes p heißt jede Menge V, die eine offene Menge O enthält, welcher der Punkt p noch angehört. $\mathfrak{V}(p)$ sei die Gesamtheit der Umgebungen von p. Eine Untermenge A eines topologischen

[1]) Die Bezeichnungen, die sich auf projektive Ebenen beziehen, lehnen sich an PICKERT [**1955**] an, die topologischen Bezeichnungen an BOURBAKI [**1951**].

Raumes P wird in natürlicher Weise zu einem topologischen Raum, wenn man die Spuren der offenen Mengen von P auf A als offene Mengen von A nimmt. Die Produktmenge zweier topologischer Räume wird mit den Produkten offener Mengen und deren Vereinigungen als offenen Mengen ein topologischer Raum. Eine Abbildung eines topologischen Raumes in einen anderen heißt stetig, wenn das volle Urbild jeder offenen Menge des Bildraumes eine offene Menge des Urbildraumes ist. Sind bei einer projektiven Ebene der Raum P der Punkte und der Raum $\mathfrak{G}$ der Geraden topologische Räume, so ist damit erklärt, was es heißt, die Bildung der Verbindungsgeraden (also die Abbildung des Raumes der Paare verschiedener Punkte auf ihre Verbindungsgeraden) sei stetig, und dual auch die Stetigkeit der Bildung des Schnittpunktes. Eine projektive Ebene soll topologisch heißen, wenn die Bildung von Verbindungsgerader und Schnittpunkt in diesem Sinne stetig ist und es nicht leere, echte, offene Teilmengen von P gibt, aber nicht alle Teilmengen von P offen sind. Die letzte Festsetzung dient nur zur Vermeidung unbequemer Sonderfälle. Ist G eine Gerade und $\mathfrak{p}$ ein nicht auf ihr gelegener Punkt, so bezeichne $G \simeq \mathfrak{p}$ die „perspektive" Abbildung $x \to x \cup \mathfrak{p}$ der Menge der Punkte auf G auf die Menge der Geraden durch $\mathfrak{p}$. Die durch Hintereinander-Ausführung von Perspektivitäten entstehenden Selbstabbildungen einer Geraden heißen Projektivitäten. Sie bilden bekanntlich eine dreifach transitive Gruppe auf der Geraden. In einer topologischen Ebene ist jede Perspektivität und damit auch jede Projektivität ein Homöomorphismus.

§ 2. Koordinaten und topologische Ternärkörper

Wählt man vier Punkte $0, e, u, v$ mit der in (III) genannten Eigenschaft und nimmt man die Gerade $u \cup v$ und alle auf ihr liegenden Punkte aus der projektiven Ebene heraus, so erhält man eine affine Ebene. Zwei Geraden der affinen Ebene, die entweder zusammenfallen oder sich nicht schneiden, heißen parallel. Durch die Festsetzung

$$(*) \qquad \begin{cases} x_{\mathfrak{p}} = \big(((\mathfrak{p} \cup v) \cap (0 \cup e)) \cup u\big) \cap (0 \cup v) \\ y_{\mathfrak{p}} = (\mathfrak{p} \cup u) \cap (0 \cup v) \end{cases}$$

werden die Punkte $\mathfrak{p}$ der affinen Ebene umkehrbar eindeutig den Paaren $(x_{\mathfrak{p}}, y_{\mathfrak{p}})$ von „Koordinaten", d.h. Elementen der Menge T der von v verschiedenen Punkte auf $0 \cup v$ zugeordnet. y_e wird mit 1 bezeichnet. Als $\tau(s, x, t)$ erklärt man dasjenige Element y, für welches (x, y) auf der Parallelen zu $(0, 0) \cup (1, s)$ durch $(0, t)$ liegt. T wird zusammen mit der ternären Verknüpfung τ zu einer als Ternärkörper bezeichneten algebraischen Struktur, in der sich alle geometrischen Beziehungen wiederfinden. Die Übersetzung der Eigenschaften (I), (II) liefert für den Ternärkörper:

(1) Zu s, x, y gibt es genau ein t mit $y = \tau(s, x, t)$.

(2) Zu $x_1 \neq x_2$ und y_1, y_2 gibt es genau ein Paar s, t mit $y_i = \tau(s, x_i, t)$ für $i = 1, 2$.

(3) Zu $s_1 \neq s_2$ und t_1, t_2 gibt es genau ein x mit $\tau(s_1, x, t_1) = \tau(s_2, x, t_2)$.

In einem Ternärkörper führt man die beiden folgenden, als Addition und Multiplikation bezeichneten binären Verknüpfungen ein:

$$x + t = \tau(1, x, t), \qquad s\,x = \tau(s, x, 0).$$

Die Menge T bildet zusammen mit der Addition eine Loop mit dem neutralen Element 0, die Menge $T^{\times} = T - 0$ zusammen mit der Multiplikation eine Loop mit dem neutralen Element 1. [In einer desarguesschen projektiven Ebene gilt $\tau(s, x, t) = s\,x + t$ und $T^{\times}$ ist zusammen mit der Addition und Multiplikation ein Schiefkörper.]

Liegt eine topologische projektive Ebene vor, so wird die affine Ebene durch die Darstellung (*) topologisches Produkt der affinen Geraden T mit sich selbst (SKORNJAKOV [**1954**], SALZMANN [**1955**]). (T, τ) wird ein topologischer Ternärkörper, das soll heißen, $\tau(s, x, t)$ hängt stetig von allen drei Veränderlichen (gleichzeitig) ab und jede der in (1), (2), (3) geforderten Umkehrungen ist stetig in allen übrigen Veränderlichen. Umgekehrt ist es aber nicht gelungen allgemein zu beweisen, daß ein topologischer Ternärkörper wieder zu einer topologischen projektiven Ebene führt. In einem topologischen Ternärkörper T bildet $T^{\times}$ zusammen mit der Addition eine topologische Loop, d.h. die Addition sowie ihre rechte und linke Umkehrung sind stetige Funktionen. Die Abbildung $x \to \tau(s, x, t)$ stellt für $s \neq 0$ einen Homöomorphismus von T dar; denn $\tau(s, x, t) = \tau(0, x, y) = y$ ist bei festem s, t $(s \neq 0)$ stetig nach x auflösbar. Die Menge dieser Abbildungen ist nach (2) wegen $\tau(0, x, t) = t$ zweifach transitiv auf T.

Ein Ternärkörper einer topologischen projektiven Ebene ist ein regulärer topologischer Raum. Hieraus folgt, daß auch jede projektive Ebene als topologischer Raum regulär ist (SKORNJAKOV [**1954**], SALZMANN [**1955**]).

§ 3. Abzählbarkeitseigenschaften in lokalkompakten projektiven Ebenen und Ternärkörpern

Ein topologischer Raum heißt kompakt, wenn sich aus jeder Überdeckung mit offenen Mengen eine Überdeckung mit endlich vielen dieser offenen Mengen auswählen läßt; er heißt lokalkompakt, wenn es zu jedem Punkt eine Umgebung gibt, die als Unterraum des Gesamtraumes kompakt ist. Man sagt, ein topologischer Raum habe eine abzählbare Basis (oder: er erfülle das zweite Abzählbarkeitsaxiom), wenn es abzählbar viele offene Mengen gibt, so daß sich jede offene Menge als Vereinigung von einigen dieser offenen Mengen darstellen läßt; man sagt, er sei lokal von abzählbarer Basis, wenn es zu jedem Punkt abzählbar viele Umgebungen V_n gibt, so daß jede Umgebung des Punktes eine Umgebung V_n enthält. In diesem Abschnitt wird gezeigt:

Jede lokalkompakte projektive Ebene ist von abzählbarer Basis.

Dazu beweist man zuerst:

Ein lokalkompakter Ternärkörper ist lokal von abzählbarer Basis.

Um das einzusehen, beachtet man zunächst, daß der Ternärkörper als nicht diskret vorausgesetzt war, also nicht alle Punkte offene Mengen sind, und wegen der Homogenität sogar kein Punkt offen ist. Berücksichtigt man noch (T_1) und die Tatsache, daß der Durchschnitt endlich vieler Umgebungen wieder eine Umgebung ist, so folgt, daß jede Umgebung unendlich viele Elemente enthält. Man betrachtet nun eine kompakte Umgebung irgendeines Punktes und in ihr eine Folge $\{a_n\}_{n=1}^{\infty}$ verschiedener Punkte a_n. Diese Folge muß auf Grund der Kompaktheit der sie enthaltenden Menge einen Häufungspunkt haben, d.h. einen solchen Punkt a, daß jede Umgebung von a wenigstens ein a_n enthält. Durch Anwendung eines Homöomorphismus kann man sich a in 0 überführt denken, so daß man also eine Folge $\{a_n\}_{n=1}^{\infty}$ finden kann, die 0 zum Häufungspunkt hat. Wählt man noch eine kompakte Umgebung W von 0, so bilden die $V_n = a_n W$ die geforderte abzählbare Umgebungsbasis von 0; denn sei V irgendeine Umgebung von 0, w ein Punkt aus W. Dann gibt es wegen $0w = 0$ eine Umgebung V_w von 0 und eine Umgebung W_w von w mit $V_w W_w \subseteq V$. Jetzt kann man W mit endlich vielen $W_{w_1}, \ldots, W_{w_k}$ der W_w überdecken. Für $V_0 = V_{w_1} \cap \cdots \cap V_{w_k}$ gilt $V_0 W \subseteq V$. Nun enthält aber V_0 ein a_n, so daß also insbesondere V_n in V enthalten ist. Infolge der Homogenität von T überträgt sich die abzählbare Umgebungsbasis auch auf jeden anderen Punkt x von T.

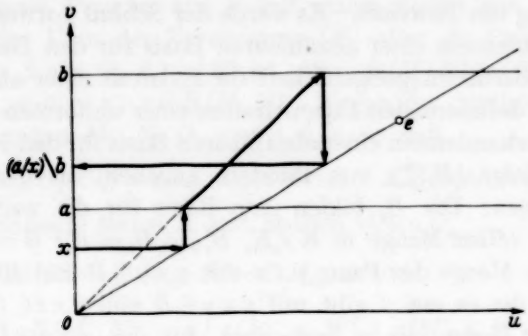
Ist der Ternärkörper einer topologischen projektiven Ebene lokalkompakt, so ist auch die projektive Ebene lokalkompakt; denn die affine Ebene ist es als Produkt eines lokalkompakten Raumes mit sich selbst, und dann ist es wegen der Abgeschlossenheit der Geraden auch die projektive Ebene.

Ein Ternärkörper einer lokalkompakten projektiven Ebene ist σ-kompakt.

Das bedeutet, er ist Vereinigung von höchstens abzählbar vielen kompakten Mengen. Beim Beweis wird wesentlich von der Einbettung des Ternärkörpers in eine topologische projektive Ebene Gebrauch gemacht. Aus einer sich gegen 0 häufenden Folge von Elementen $\neq 0$ werde eine gegen 0 konvergente Teilfolge $\{a_n\}_{n=1}^{\infty}$ ausgewählt, $\lim_{n\to\infty} a_n = 0$. Das ist möglich, da der Raum lokal von abzählbarer Basis ist. Ferner sei a ein von 0 verschiedenes Element. Nach der multiplikativen Loopeigenschaft von $T^\times$ bestimmt man s_n mit $s_n a_n = a$. Wählt man nun noch eine kompakte Umgebung W von 0, so ist jedes Element b von T in einem $s_n W$ enthalten, d.h. aber nichts anderes als $T = \bigcup_{n=1}^{\infty} s_n W$, und jedes $s_n W$ ist homöomorph zu W, also kompakt. Schreibt man nämlich für $xy = z$ auch $x = z/y$ oder $y = x\backslash z$ und setzt man $b = s_n b_n$, wobei man sich auf $b \neq 0$ beschränken kann, so ist $b_n = (a/a_n)\backslash b$ und es bleibt noch zu zeigen, daß mit a_n auch die b_n gegen 0 konvergieren, also schließlich in W liegen. Hierzu muß man auf die geometrische Definition der Multiplikation zurückgehen. Die Abbildung $x \to (a/x)\backslash b$ wird für $a \neq 0$, $b \neq 0$ durch eine Projektivität der Geraden $0 \vee v$ gegeben, und zwar bildet man $0 \vee v$ perspektiv auf das Geradenbüschel in u ab, $0 \vee v \simeq u$ und weiter

$$u \simeq 0 \vee e \simeq v \simeq a \vee u \simeq 0 \simeq b \vee u \simeq v \simeq 0 \vee e \simeq u \simeq 0 \vee v.$$

Sie stellt also einen Homöomorphismus von $0 \cup v$ dar. Dabei geht, wie man nachrechnet, 0 in sich über, also jede gegen 0 konvergente Folge wieder in eine solche.



Aus den beiden eben bewiesenen Sätzen folgt:

Eine lokalkompakte projektive Ebene ist lokal von abzählbarer Basis und σ-kompakt.

Denn beide Eigenschaften übertragen sich von den Faktoren auf das Produkt zweier topologischer Räume, also von dem Ternärkörper auf die affine Ebene. Die abzählbare Umgebungsbasis ist dann natürlich auch in der projektiven Ebene vorhanden. Ferner ist die projektive Gerade σ-kompakt, also auch die projektive Ebene als Vereinigung zweier σ-kompakter Räume, nämlich der affinen Ebene und der „uneigentlichen" Geraden $u \cup v$.

Jetzt wird die Existenz einer abzählbaren Basis gezeigt. Der Beweis wird zuerst für eine kompakte Teilmenge K eines zu der Ebene gehörigen Ternärkörpers T geführt. Er beruht auf den folgenden Gedanken: Bei einem kompakten Raum K erzeugen die Umgebungen der Diagonale D von $K \times K$ [die offenen Mengen von $K \times K$, welche die Menge D der Paare $x \times x$ $(x \in K)$ enthalten] eine uniforme Struktur, und zwar die einzige, welche mit der Topologie von K verträglich ist (Bourbaki [**1951**], Top. gén. II § 4, Th. 1).

Diese Umgebungen von D sollen als „Bänder" bezeichnet werden. Gibt es nun eine abzählbare Basis für den Filter der Bänder und gilt außerdem für den topologischen Raum die Lindelöfsche Überdeckungseigenschaft, die besagt, daß sich aus jeder Überdeckung mit offenen Mengen abzählbar viele dieser offenen Mengen auswählen lassen, die den Raum auch schon überdecken, so hat der Raum eine abzählbare Basis. Es wird sich nun zeigen, daß man die Bänder der uniformen Struktur von K auf die folgende Weise erhalten kann: Man nimmt die Menge aller Punkte, die auf irgendeiner von den Geraden liegen, die zu einer Umgebung der Geraden $0 \cup e$ (in der Topologie des Raumes $\mathfrak{G}$ aller Geraden) gehören und schränkt diese Menge auf $K \times K$ ein. Nun zeigt man zunächst, daß zugleich mit P auch der Raum $\mathfrak{G}$ lokal von abzählbarer Basis ist. Damit hat man dann eine abzählbare Basis für den Bandfilter der uniformen Struktur von K gefunden. Die Lindelöfsche Überdeckungseigenschaft ergibt sich als Abschwächung der Kompaktheit

von K. Mit Hilfe der σ-Kompaktheit von T kann man dann eine abzählbare Basis für T angeben. Die Konstruktion einer solchen für P macht hierauf keine Schwierigkeiten mehr.

Durchführung des Beweises. Es werde der Schluß vorweggenommen, daß aus dem Vorhandensein einer abzählbaren Basis für den Bandfilter und der Lindelöfschen Überdeckungseigenschaft die Existenz einer abzählbaren Basis folgt. Nach den definierenden Eigenschaften einer uniformen Struktur von K kann man bei Vorhandensein einer abzählbaren Basis für den Filter der Bänder induktiv eine Folge $\{B_i\}_{i=1}^{\infty}$ von Bändern angeben, die den folgenden Bedingungen genügen: Die B_i bilden eine Basis für die uniforme Struktur, jedes B_i ist eine offene Menge in $K\times K$, $B_{i+1}^2 \subseteqq B_i = B_i^{-1}$ $(i=1, 2, \ldots)$. Dabei bedeutet B^{-1} die Menge der Paare $y\times x$ mit $x\times y\in B$ und B^2 die Menge der Paare $x\times z$, für die es ein y gibt mit $x\times y\in B$ und $y\times z\in B$. Ferner werde mit $V(x, B)$ die Menge der y bezeichnet, für die $x\times y\in B$. Die $V(x, B_i)$ bilden dann eine Umgebungsbasis für den Punkt x. Läßt man x alle Punkte des Raumes K durchlaufen, so erhält man eine Basis für das System der offenen Mengen von K.

Hieraus muß nun eine passende abzählbare Teilmenge ausgesucht werden, die auch noch eine Basis für die offenen Mengen bildet. Nach der Lindelöfschen Überdeckungseigenschaft kann man K für jedes feste i mit abzählbar vielen $V(x_{ik}, B_{i+1})$ überdecken. Wenn man jetzt zeigen kann, daß jedes $V(x, B_i)$ ein $V(x_{ik}, B_{i+1})$ enthält, das seinerseits noch Umgebung von x ist, dann ist damit bewiesen, daß die Gesamtheit der $V(x_{ik}, B_{i+1})$ $(i, k=1, 2, \ldots)$ eine abzählbare Basis der gewünschten Art ist. Da nun die $V(x_{ik}, B_{i+1})$ den ganzen Raum überdecken, gibt es ein x_{ik} mit $x\in V(x_{ik}, B_{i+1})$, d.h. $x_{ik}\times x\in B_{i+1}$ oder $x\times x_{ik}\in B_{i+1}$. Gilt jetzt noch $x_{ik}\times y\in B_{i+1}$, also $y\in V(x_{ik}, B_{i+1})$, so folgt $x\times y\in B_i$, $y\in V(x, B_i)$, und daher $x\in V(x_{ik}, B_{i+1}) \subseteqq V(x, B_i)$.

Um nun zu zeigen, daß die Bänder von K in der angegebenen Weise erhalten werden können, geht man so vor: Es sei $B(M, N)$ die Menge der Paare $x\times y\in K\times K$ mit $y\in\tau(M, x, N)$, wo $\tau(M, X, N)$ entsprechend der Komplexmultiplikation in der Gruppentheorie zu verstehen ist. Läßt man jetzt M eine abzählbare Basis von Umgebungen der 1 in der Topologie von T durchlaufen und N eine abzählbare Basis von Umgebuugen der 0 (man kann sich hierbei auf offene Mengen M, N beschränken), so erhält man in der Gesamtheit der $B(M, N)$ eine abzählbare Basis für den Filter der Bänder von K. Dazu muß man nur noch nachweisen, daß die $B(M, N)$ den Bandfilter erzeugen, der zu der Topologie von K gehört. Zum Nachweis der Bandfiltereigenschaft ist nur zu zeigen, daß es zu gegebenem M, N Mengen M', N' mit $B^{-1}(M', N')\circ B(M', N') \subseteqq B(M, N)$ gibt. B^{-1} ist wie vorher erklärt, $B\circ C$ ist das relationstheoretische Produkt von B und C, also die Menge der $x\times z$, für die es ein y gibt mit $x\times y\in B$ und $y\times z\in C$. Sind M, N offene Mengen von T, so ist $B(M, N)$ eine offene Menge von $K\times K$, wie sich folgendermaßen ergibt: Die Menge der Geraden der Form $y=\tau(m, x, n)$ mit $m\in M$, $n\in N$ bildet eine offene Umgebung $\mathfrak{S}$ von $0\smile e$. Betrachtet man eine solche Gerade

als Verbindungsgerade von (x, y) mit einem zweiten Punkt $(x', \tau(m, x', n))$ auf ihr, so folgt aus den Stetigkeitsforderungen, daß es eine Umgebung V von (x, y) und eine Umgebung V' des zweiten Punktes so gibt, daß die Verbindungsgerade eines Punktes aus V mit einem Punkt aus V' stets zu $\mathfrak{S}$ gehört. Dann liegt V in der Vereinigung $[\mathfrak{S}]$ aller als Punktmengen aufgefaßten Geraden S aus $\mathfrak{S}$, d.h. $[\mathfrak{S}]$ ist offen, $B(M, N) = [\mathfrak{S}] \cap K \times K$. Für $x \in K$ ist $x \times \tau(1, x, 0) = x \times x \in B(M, N)$. Wegen der Stetigkeit von τ gibt es offene Umgebungen M_x, N_x, O_x von $1, 0, x$, so daß

$$\tau(M_x, O_x, N_x) \times \tau(M_x, O_x, N_x) \cap K \times K \subseteq B(M, N).$$

Man kann die kompakte Menge K mit endlich vielen $O_{x_1}, \ldots, O_{x_n}$ der O_x überdecken. Setzt man $M' = \bigcap_{i=1}^{n} M_{x_i}$, $N' = \bigcap_{i=1}^{n} N_{x_i}$, so ist

$$B^{-1}(M', N') \circ B(M', N') = \bigcup_{x \in K} \tau(M', x, N') \times \tau(M', x, N') \cap K \times K,$$

also wirklich in $B(M, N)$ enthalten. Für offene M, N bilden daher die $B(M, N)$ eine Basis von offenen Bändern, erzeugen also auf K eine Topologie, die höchstens gröber ist, als die auf K vorgegebene. Nun ist aber der Durchschnitt der $B(M, N)$, wenn man M, N das System der Umgebungen von $1, 0$ durchlaufen läßt, nichts anderes als die Diagonale D von K (wegen der Stetigkeit von τ). Daher ist die von den $B(M, N)$ erzeugte Topologie regulär. Nun kann eine kompakte Topologie nicht echt feiner sein, als irgendeine andere reguläre Topologie auf demselben Raum; denn ist E' ein Hausdorff-Raum, E dieselbe Punktmenge mit einer kompakten Topologie, die feiner oder gleich der von E' ist, dann ist die identische Abbildung von E auf E' stetig, führt also die abgeschlossenen Mengen von E, die ja in E kompakt sind, in kompakte Mengen von E' über, und kompakte Mengen eines Hausdorff-Raumes sind wieder abgeschlossen: Die Topologie von E' ist auch feiner oder gleich der von E, d.h. die beiden Topologien sind gleich. Also muß die von den $B(M, N)$ erzeugte Topologie mit der gegebenen auf K übereinstimmen.

Hiermit ist bewiesen: Jede kompakte Teilmenge K von T hat eine abzählbare Basis für die Topologie. Nun stellt man, was nach dem Beweis für die σ-Kompaktheit von T möglich ist, T dar als Vereinigung von abzählbar vielen offenen Mengen Q_n, so daß jedes Q_n in einer kompakten Menge enthalten ist. Dann hat jedes Q_n eine abzählbare Basis. Nimmt man alle diese zusammen, so erhält man eine abzählbare Basis von T. Mit T hat auch $T \times T$ und damit jede affine Ebene eine abzählbare Basis. Überdeckt man jetzt P mit drei offenen Mengen, von denen jede eine ganze Gerade nicht enthält, und nimmt man die abzählbaren Basen für diese offenen Mengen zusammen, so erhält man schließlich auch die abzählbare Basis für P.

§ 4. Kompaktheit der projektiven Ebene

Die reelle affine Ebene ist lokalkompakt, die reelle projektive Ebene ist sogar kompakt. Das legt die Frage nahe, ob oder unter welchen Voraussetzungen aus der lokalen Kompaktheit einer topologischen projektiven Ebene

ihre Kompaktheit folgt. Ein erstes Ergebnis in dieser Richtung, das wegen seiner schlechten Zugänglichkeit hier mit Beweis wiedergegeben werden soll, stammt von SKORNJAKOV [**1954**]:

Eine topologische projektive Ebene, deren Geraden kompakt sind, ist selbst kompakt.

Damit ist die gestellte Frage darauf zurückgeführt, wann die Geraden einer lokalkompakten Ebene kompakt sind; denn sind die Geraden kompakt, so ist die Ebene lokalkompakt, und ist die projektive Ebene kompakt, so auch die projektiven Geraden als abgeschlossene Punktmengen in der Ebene. Der Hauptschritt des Beweises ist nun der folgende: Sind die Geraden einer projektiven Ebene kompakt und ist $\mathfrak{L}$ eine kompakte Menge im Raum $\mathfrak{G}$ aller Geraden, so daß die Menge $[\mathfrak{L}]$ der Punkte, die auf irgendeiner Geraden aus $\mathfrak{L}$ liegen (d.h. die Vereinigung aller als Punktmengen aufgefaßten Geraden L aus $\mathfrak{L}$) nicht der ganze Raum P ist, so ist $[\mathfrak{L}]$ kompakt. Um das einzusehen, bedient sich SKORNJAKOV eines Kompaktheitskriteriums, das überabzählbare absteigende Folgen abgeschlossener Mengen verwendet. Wegen der bewiesenen Existenz einer abzählbaren Basis kommt man aber mit abzählbaren Folgen aus und kann das übliche Kriterium benutzen. Um für eine Folge nicht leerer abgeschlossener Mengen $A_n \subseteq [\mathfrak{L}]$ zu beweisen, daß sie nicht leeren Durchschnitt hat, wählt man einen Punkt p, der nicht in $[\mathfrak{L}]$ liegt und eine nicht durch p gehende Gerade G, und zeigt zunächst, daß jede abgeschlossene Menge $A \subseteq [\mathfrak{L}]$ bei der Projektion aus p auf G, d.h. der Abbildung $x \to (x \cup p) \cap G$ von A auf eine Menge B in G, wieder in eine abgeschlossene Menge übergeht. Hierbei geht wesentlich die Kompaktheit von $\mathfrak{L}$ ein, und zwar so: Jedem Punkt x aus B ordnet man irgendeine Gerade L_x aus $\mathfrak{L}$ zu, die durch irgendeinen Urbildpunkt von x hindurchgeht. Diese Auswahl der L_x werde jetzt festgehalten. Häuft sich eine Menge von Punkten x aus B gegen einen Punkt z, der dann natürlich auch auf G liegt, so kann man eine Gerade L aus $\mathfrak{L}$ finden, gegen die sich die entsprechenden L_x häufen, genauer gesagt, eine Gerade L, so daß es in jeder beliebigen Umgebung von L Geraden L_x mit x beliebig nahe bei z gibt. Wäre das nämlich nicht der Fall, so gäbe es zu jeder Geraden L aus $\mathfrak{L}$ eine Umgebung $\mathfrak{S}$ und dazu eine Umgebung V von z, so daß $L_x \notin \mathfrak{S}$ für alle $x \in V \cap B$ gilt. Nun könnte $\mathfrak{L}$ mit endlich vielen der $\mathfrak{S}$ überdeckt werden, und dann wäre der Durchschnitt der entsprechenden endlich vielen V frei von Punkten x aus B im Widerspruch dazu, daß sich die x gegen z häufen. Für die so gefundene Gerade L ist $y = L \cap (p \cup z)$ Häufungspunkt von A; denn in einer beliebigen Umgebung von y findet sich ein Punkt $L_x \cap (p \cup x)$, d.h. ein nach Definition zu A gehöriger Urbildpunkt von x. Da A abgeschlossen war, gehört y zu A. Bei der Projektion geht y in den Punkt z über, der somit zu B gehört, d.h. B ist abgeschlossen. Die A_n gehen also bei der Projektion in eine absteigende Folge nicht leerer abgeschlossener Mengen $B_n \subseteq G$ über. Da G kompakt ist, haben alle B_n wenigstens einen Punkt q gemeinsam. Die Gerade $H = p \cup q$ hat nun mit allen A_n einen nicht leeren Durchschnitt. Da auch H kompakt ist, muß der Durchschnitt der A_n mit H noch Punkte gemeinsam haben,

und damit ist die Behauptung über die Kompaktheit von $[\mathfrak{L}]$ bewiesen. Um jetzt zur Kompaktheit von ganz P zu gelangen, sucht man P zu überdecken mit einer Menge $[\mathfrak{L}]$ der eben betrachteten Art und einer kompakten Umgebung eines nicht zu $[\mathfrak{L}]$ gehörigen Punktes. Das gelingt auf folgende Weise: Es seien p, p_1, p_2 nicht kollineare Punkte. Auf der Geraden $p \cup p_i$ wählt man nun kompakte Mengen U_i, V_i, deren Vereinigung die ganze Gerade $p \cup p_i$ selbst ist, und zwar so, daß $p_i \notin U_i$ und $p \notin V_i$. Dann ist die zu $U_1 \times U_2$ homöomorphe Menge $K = (U_1 \cup p_2) \cap (U_2 \cup p_1)$ kompakt. Ebenso ist die zu $V_1 \times V_2$ homöomorphe Geradenmenge $V_1 \cup V_2$ eine kompakte Menge in $\mathfrak{G}$. Nach dem ersten Beweisschritt ist also $[V_1 \cup V_2]$ kompakt. Wegen $p_i \in V_i$ überdeckt $[V_1 \cup V_2]$ aber $P - K$. Daher ist P als Vereinigung von zwei kompakten Mengen kompakt.

SKORNJAKOV schließt nun irrtümlicherweise, daß man die Geraden einer lokalkompakten topologischen projektiven Ebene durch Projektion einer kompakten Umgebung eines Punktes aus diesem Punkt als Projektionszentrum als stetiges Bild einer kompakten Menge erhalten könne, und daß sie somit selbst kompakt seien. Dabei übersieht er, daß die Projektion im Zentrum selbst nicht erklärt ist und dort auch nicht stetig ergänzt werden kann, und daß die um das Zentrum verminderte Umgebung im allgemeinen nicht mehr kompakt sein wird. Das Versagen seiner Schlußweise zeigt sich schon an der diskreten Ebene, die SKORNJAKOV nicht ausschließt. In ihr ist jeder Punkt eine kompakte Umgebung von sich selbst; also ist sie lokalkompakt. Kompakt ist sie aber nur, wenn sie endlich ist. Weitere Beispiele lokalkompakter, aber nicht kompakter topologischer projektiver Ebenen sind nicht bekannt.

Es gibt nun aber einige, in ihrem Wesen ganz verschiedene, Zusatzvoraussetzungen, unter denen der Schluß von der lokalen Kompaktheit der Ebene auf die „globale" Kompaktheit gezogen werden kann. Die erste ist rein topologischer Art und besagt den Zusammenhang der topologischen Ebene; die zweite setzt die topologische Struktur der Ebene in nähere Beziehung zur geometrischen, und zwar soll die Topologie durch die Anordnung der Ebene hervorgerufen werden; die dritte schließlich ist von rein geometrischer, oder man kann auch sagen, algebraischer Natur, indem sie gewisse Spezialfälle des Schließungssatzes von DESARGUES fordert, die bei Übersetzung in die Sprache des Ternärkörpers zu bestimmten schwachen Assoziativitätsgesetzen für die Multiplikation führen. Auf alle diese Dinge wird an späteren Stellen eingegangen werden.

§ 5. Zusammenhangsverhältnisse in lokalkompakten Ebenen

Ein topologischer Raum heißt zusammenhängend, wenn er nicht in die Summe zweier nicht leerer offener Mengen zerlegt werden kann. (Dabei soll unter „Summe" die Vereinigung paarweise disjunkter Mengen verstanden werden.) Da das Komplement einer offenen Menge abgeschlossen ist, bedeutet Zusammenhang nichts anderes, als daß es in dem Raum keine zugleich offenen und abgeschlossenen Mengen gibt außer der leeren Menge und dem

ganzen Raum selbst. Unter der zusammenhängenden Komponente eines Punktes x in einem topologischen Raum versteht man die größte, den Punkt x enthaltende zusammenhängende Menge in dem Raum. Sie ergibt sich als Vereinigung aller den Punkt enthaltenden zusammenhängenden Mengen und ist immer vorhanden. Ein Raum, in dem die zusammenhängende Komponente jedes Punktes nur aus dem Punkt selbst besteht, heißt völlig unzusammenhängend. Als den Rand $rd\,A$ einer Menge A bezeichnet man die Differenz zwischen der kleinsten abgeschlossenen Menge, welche A umfaßt, und der größten offenen Menge, die in A enthalten ist. Genau die zugleich offenen und abgeschlossenen Mengen haben also keinen Rand. Hat ein Punkt eine randlose Umgebung, so muß die Komponente des Punktes in dieser enthalten sein; denn sonst ergäbe sich eine Zerlegung der Komponente in die Summe zweier abgeschlossener Mengen. Der Durchschnitt aller randlosen Umgebungen eines Punktes x heißt die Quasikomponente von x und umfaßt nach dem eben Gesagten die zusammenhängende Komponente von x. Bestehen alle Quasikomponenten in einem Raum nur aus einzelnen Punkten, so heißt der Raum nirgends zusammenhängend. Gibt es sogar zu jedem Punkt beliebig kleine randlose Umgebungen, mit anderen Worten, bilden die randlosen Umgebungen eines Punktes eine Basis für seinen Umgebungsfilter, so heißt der Raum nulldimensional.

Für beliebige topologische projektive Ebenen war früher bewiesen worden (SALZMANN [**1955**]):

Nimmt man aus einer zusammenhängenden projektiven Ebene einen Punkt oder eine ganze Gerade heraus, so ist die verbleibende Punktmenge auch noch zusammenhängend. Die Geraden einer zusammenhängenden projektiven oder affinen Ebene sind zusammenhängend. Ist die projektive Ebene nicht zusammenhängend, so ist sie sogar nirgends zusammenhängend (und mit ihr natürlich auch alle Untermengen).

Für lokalkompakte projektive Ebenen lassen sich diese Aussagen wesentlich einfacher beweisen, als im allgemeinen Fall, und auch noch verschärfen. Die wichtigste Verschärfung ist die, daß eine lokalkompakte zusammenhängende projektive Ebene auch lokalzusammenhängend ist; das bedeutet, daß es zu jedem Punkt beliebig kleine zusammenhängende Umgebungen gibt. Zuerst ergibt sich anknüpfend an den Beweis für die Existenz einer abzählbaren Umgebungsbasis in einem lokalkompakten Ternärkörper:

Ein lokalkompakter topologischer Ternärkörper (und insbesondere eine Gerade einer lokalkompakten affinen Ebene) *ist entweder nulldimensional oder zusammenhängend und lokalzusammenhängend.*

Die Homöomorphismen eines topologischen Raumes, die einen Punkt x festlassen, führen nämlich auch die zusammenhängende Komponente von x in sich über. Bei einem topologischen Ternärkörper bilden diese Homöomorphismen aber wegen der zweifachen Transitivität der Gruppe aller Homöomorphismen eine Gruppe, die auf den von x verschiedenen Punkten noch

einfach transitiv ist. Enthält also die zusammenhängende Komponente eines Punktes noch einen weiteren Punkt, so enthält sie sogar alle Punkte. Daher ist ein topologischer Ternärkörper entweder völlig unzusammenhängend oder zusammenhängend. Man kann diesen selben Schluß auch mit der Quasikomponente durchführen und erhält dann etwas stärker: Ein topologischer Ternärkörper ist entweder nirgends zusammenhängend oder zusammenhängend. Nun sei W eine kompakte Umgebung von 0. Im ersten Fall besteht dann die Quasikomponente von 0 in dem kompakten Raum W erst recht nur aus dem Element 0 allein. Um hieraus die Nulldimensionalität zu beweisen, muß man zu einer beliebig kleinen, ganz in W enthaltenen, offenen Umgebung U von 0 eine randlose Umgebung von 0 angeben können, die in U enthalten ist. Die Menge $W - U$ ist abgeschlossen und daher kompakt. Für jeden Punkt x aus $W - U$ kann man, da ja der Durchschnitt aller randlosen Umgebungen von 0 nur aus dem Element 0 selbst besteht, eine randlose Umgebung V_x von 0 finden, die den Punkt x nicht enthält. $W - V_x$ ist eine Umgebung von x in dem Raum W. Die kompakte Menge $W - U$ kann mit endlich vielen $W - V_{x_1}, \ldots, W - V_{x_n}$ dieser Umgebungen überdeckt werden, dann ist der Durchschnitt $\bigcap_{i=1}^{n} V_{x_i}$ in U enthalten, und er ist eine randlose Umgebung von 0.

Im zweiten Fall, wo der Ternärkörper zusammenhängend ist, kann also auch die in einer kompakten Umgebung W von 0 gebildete Quasikomponente C von 0 nicht nur aus dem Element 0 allein bestehen. C ist als Durchschnitt abgeschlossener Mengen abgeschlossen und daher kompakt. Darum gibt es, wie in § 3 bewiesen, eine Umgebung V von 0 mit $VC \subseteq W$. (Hierbei ist VC das Komplexprodukt in der multiplikativen Loop des Ternärkörpers.) Sieht man VC an als die Vereinigung der Mengen vC, wo v die Menge V durchläuft, so zeigt sich, daß VC als Vereinigung von zusammenhängenden Mengen mit paarweise nicht leerem Durchschnitt (0 ist enthalten in vC) selbst zusammenhängend ist; sieht man VC an als die Vereinigung der Mengen Vc, wo c die Menge C durchläuft, so ergibt sich weiter, daß VC eine ganze Umgebung von 0 enthält und damit selbst Umgebung von 0 ist. Die abgeschlossene Hülle $\overline{VC}$ von VC ist zusammenhängend und als Untermenge von W kompakt. Für eine gegen 0 konvergente Folge a_n bilden also die $a_n \overline{VC}$, wie in § 3 bewiesen, eine Basis von Umgebungen der 0. Jede Menge $a_n \overline{VC}$ ist homöomorph zu $\overline{VC}$ und somit zusammenhängend. Für jeden anderen Punkt des Ternärkörpers gilt Entsprechendes. Damit ist in diesem Fall der lokale Zusammenhang des Ternärkörpers bewiesen. Ist nun eine Gerade einer affinen Ebene nulldimensional, so auch die affine Ebene selbst als topologisches Produkt zweier Geraden, aber dann ist auch die projektive Ebene nulldimensional und folglich nicht zusammenhängend. Umgekehrt überträgt sich auch Zusammenhang und lokaler Zusammenhang von der affinen Geraden auf die affine Ebene und weiter auf die projektive Ebene, die ja die abgeschlossene Hülle der affinen Ebene ist. (Die abgeschlossene Hülle einer zusammenhängenden Menge ist immer zusammenhängend.) Damit hat man:

Eine lokalkompakte projektive Ebene ist entweder nulldimensional oder zusammenhängend und lokalzusammenhängend. Diese Eigenschaften kommen gleichzeitig mit der Ebene auch ihren Geraden zu.

Nun ergibt sich ganz einfach:

Eine zusammenhängende lokalkompakte projektive Ebene ist kompakt.

Nach dem früher Bewiesenen hat man dazu nur noch zu zeigen: Die Geraden einer zusammenhängenden lokalkompakten projektiven Ebene sind kompakt. Dazu betrachtet man eine kompakte Umgebung V eines Punktes p, die eine ganze Gerade G nicht trifft. Der Rand rdV von V ist als abgeschlossene Teilmenge von V kompakt, und er enthält nicht den Punkt p. Nun wird behauptet: Projiziert man rdV aus p auf die (nicht durch p gehende) Gerade G, so erhält man als Bild die ganze Gerade G. Es sei x ein beliebiger Punkt von G; dann muß also gezeigt werden, daß die Gerade $p \cup x$ mit rdV wenigstens einen Punkt gemeinsam hat, der bei der Projektion in x übergeht. Nun schneiden die „ebenen" Umgebungen von p aus einer Geraden durch p die „linearen" Umgebungen von p in der Topologie der Geraden aus. Hätte der Rand von V mit $p \cup x$ keinen Punkt gemeinsam, so bedeutete das, daß es in der linearen Topologie auf $p \cup x$ eine randlose Umgebung von p gäbe, die aber nicht die ganze Gerade $p \cup x$ wäre. Dann wäre aber $p \cup x$ nicht zusammenhängend im Widerspruch zu dem vorherigen Satz. Also ist G und damit jede Gerade als stetiges Bild einer kompakten Menge kompakt.

§ 6. Kompaktheit angeordneter Ebenen

Eine Ebene heißt angeordnet, wenn auf ihren Geraden eine Trennbeziehung erklärt ist, die bei Projektionen erhalten bleibt. Dabei ist eine Trennbeziehung eine symmetrische Relation $ab|cd$ zwischen Paaren a, b und c, d von Punkten auf einer Geraden, die erklärt ist, wenn je zwei der vier Punkte verschieden sind, und die den folgenden Bedingungen genügt: Es gilt genau eine der drei Beziehungen $ab|cd$, $bc|ad$, $ca|bd$; aus $ab|cd$ folgt $ba|cd$; aus $ab|cd$ und $bc|de$ folgt $cd|ea$. In einer angeordneten projektiven Ebene macht man einen Ternärkörper zu einer angeordneten Menge, indem man, mit der Bezeichnung ∞ für den Punkt v, festsetzt:

$a < 0$ genau für $1a|0\infty$; $0 < a$ für $a \neq 0$ sonst;

falls $0 < a$, sei $a < b$ genau für $0b|a\infty$; $b < a$ für $b \neq a$ sonst;

falls $a < 0$, sei $b < a$ genau für $0b|a\infty$; $a < b$ für $b \neq a$ sonst.

Eine geordnete Menge wird zu einem topologischen Raum, wenn man als Umgebungen eines Punktes x die offenen Intervalle (a, b) nimmt, die den Punkt x enthalten, also die Mengen der z mit $a < z < b$, wenn $a < x < b$. Führt man in einer angeordneten projektiven Ebene auf diese Weise eine Topologie ein, so wird die Ebene dadurch zu einer topologischen projektiven Ebene im hier gebrauchten Sinne. Dazu muß man über das bei PICKERT [**1955**], S. 269 Bewiesene hinaus nur noch zeigen, daß es sich nicht um die diskrete

Topologie handeln kann. Wäre nämlich der Ternärkörper durch die Anordnung diskret topologisiert, so gäbe es ein Intervall (a, b), das nur das Element 0 enthielte. Das kann aber nicht sein; denn (s. z.B. PICKERT [**1955**], S. 234) es gibt in einer angeordneten Ebene zwischen je zwei Punkten einen (und damit sogar unendlich viele) weitere.

Mit Hilfe der Invarianz der Trennbeziehung gegenüber Projektionen kann man Monotoniegesetze für die multiplikative Loop herleiten und zwar gilt $ac<bc$ und $ca<cb$ genau dann, wenn $0<c$ und $a<b$ oder wenn $c<0$ und $b<a$ ist. Beweis. Die Multiplikation von rechts- bzw. von links mit einem festen Element $w \neq 0, \infty$ ist eine Projektivität der Geraden $0 \cup v$ und läßt also die Trennbeziehung ungeändert, d.h. $xy|z\infty$ ist gleichwertig mit $xw\ yw|zw\infty$ und ebenso mit $wx\ wy|wz\infty$. Im Fall $c<0<a<b$ gilt $1c|0\infty$ und $0b|a\infty$. Hieraus folgt $a\,ac|0\infty$ und $0\,bc|ac\infty$. Nun ist aber (s. z.B. PICKERT [**1955**], S. 226) $pq|0\infty$ dasselbe wie $p<0<q$ oder $q<0<p$. Wegen $0<a$ ist also $ac<0$ und damit $bc<ac$. Die anderen Fälle erledigen sich vollkommen analog, da die Negation der Trennbeziehung ebenso wie diese selbst invariant gegenüber Projektivitäten ist.

Für einen Ternärkörper einer topologischen projektiven Ebene ist die Abbildung $x \to x' = x \backslash 1$, ergänzt durch die Festsetzungen $0' = \infty\ (= v)$, $\infty' = 0$, als die durch die Hintereinanderausführung der Projektionen der Geraden $0 \cup v$ aus u auf die Gerade $e \cup v$, dieser aus 0 auf die Gerade $e \cup u$, weiter aus v auf die Gerade $0 \cup e$ und schließlich aus u zurück auf $0 \cup v$ entstehende Projektivität ein Homöomorphismus von $0 \cup v$. Man kann also die Umgebungen von ∞ mit Hilfe dieser Abbildung aus denen von 0 gewinnen. Um nun zu beweisen, daß aus der lokalen Kompaktheit der angeordneten Ebene (in der Ordnungstopologie) ihre Kompaktheit folgt, konstruiert man unter Ausnutzung der Monotonieeigenschaften der Multiplikation eine kompakte Umgebung von ∞ und eine solche von 0, die zusammen den um den Punkt ∞ vermehrten Ternärkörper überdecken. Dazu nimmt man eine Umgebung (a, b) von 0, deren abgeschlossene Hülle kompakt ist. Diese abgeschlossene Hülle besteht aus dem abgeschlossenen Intervall $[a, b]$, d.h. aus (a, b) und den hinzugenommenen Endpunkten a, b, da ja nach dem vorher Bemerkten der Ternärkörper dicht geordnet ist, also zwischen je zwei Elementen noch weitere liegen. Mit $[a, b]$ ist aber auch jedes Intervall $[ca, cb]$ für $0<c$ eine kompakte Umgebung von 0. Andererseits geht $[a, b]$ bei der Abbildung $x \to x'$ in eine kompakte Umgebung von ∞ über. Aus $0<x<b$ wird dabei $b'<x'$; denn x' ist definiert durch $xx' = 1$, deswegen bedeutet $b'<x'$ dasselbe wie $xb'<1=bb'$. Dabei ist $0<b'$; denn $b'<0$ hätte $bb'=1<0$ zur Folge. Also ist $xb'<1$ weiter gleichwertig mit $x<b$. Entsprechend wird $a<x<0$ zu $x'<a'$. Es gibt also eine kompakte Umgebung von ∞ in der Form der Menge aller z mit $z \leqq a'$ oder $b' \leqq z$. Es gilt, nun noch ein Element c zu finden, so daß $(a', b') \subseteq [ca, cb]$. Dazu wählt man zuerst ein Hilfselement d mit $0<d$, so daß da gleich einem beliebigen Element $h<a'$ wird und dann c mit $d<c$, so daß $b'<cb$. Damit hat man die Kompaktheit der Geraden $0 \cup v$ bewiesen und es hat sich ergeben:

Ist eine angeordnete projektive Ebene lokalkompakt in der Ordnungstopologie, so ist sie kompakt.

Weiter folgt in diesem Fall noch:

Ist eine angeordnete projektive Ebene lokalkompakt in der Ordnungstopologie, so ist sie archimedisch angeordnet.

Hierfür ist zu zeigen, daß die additive Loop jedes Ternärkörpers der Ebene archimedisch angeordnet ist, daß also für beliebige Elemente a, b aus $0<a<b$ oder $b<a<0$ die Existenz einer natürlichen Zahl n folgt, für die $b<a \cdot n$ bzw. $a \cdot n<b$ gilt, wobei $a \cdot n$ rekursiv definiert ist durch $a \cdot 1=a$, $a \cdot (n+1)=a \cdot n+a$. Im ersten Fall $0<a<b$ ergibt sich die Existenz einer solchen Zahl n so: Wäre niemals $b<a \cdot n$, so wäre $a \cdot n \in [0, b]$ für jedes n. Nun ist $[0, b]$ aber kompakt, wie vorher für ein beliebiges abgeschlossenes Intervall gezeigt wurde. Also hätte die Folge $a \cdot n$ einen Häufungspunkt c in $[0, b]$. Außerdem ist die Folge $a \cdot n$ (wegen der Monotoniegesetze der Addition) monoton, und daher hätte man dann sogar die Konvergenz $\lim\limits_{n\to\infty} a \cdot n=c$. Daraus folgte $\lim\limits_{n\to\infty} a \cdot n+a=\lim\limits_{n\to\infty} a \cdot (n+1)=c+a$. Die Folgen $a \cdot n$ und $a \cdot (n+1)$ müssen aber den gleichen Grenzwert haben. Dann müßte $c=c+a$ sein und $a=0$. Das ist ein Widerspruch. Im zweiten Fall schließt man analog.

Im bisherigen ersten Teil der Arbeit wurden an den Ternärkörper nur topologische oder allenfalls ordnungstheoretische Forderungen gestellt, aber keinerlei besondere algebraische Rechenregeln für die ternäre Verknüpfung verlangt, die sich geometrisch durch gewisse Schließungssätze ausdrücken. Das soll nun im zweiten Teil der Arbeit geschehen, und zwar werden diese geometrischen Schließungssätze der einfacheren Formulierung halber meist in ihrer algebraischen Form für einen Ternärkörper ausgesprochen. Die Übersetzung in die geometrische Redeweise findet man bei PICKERT [**1955**]. Man kann dann einmal fragen, inwieweit sich bei hinreichender „Anständigkeit" des Ternärkörpers weitere Folgerungsbeziehungen zwischen topologischen Eigenschaften der Ebene ergeben. Ein Beispiel hierfür bietet die Frage nach der Kompaktheit lokalkompakter Ebenen und nach ihren möglichen topologischen Dimensionen. Zum anderen kann man auch unter topologischen Voraussetzungen nach engeren Abhängigkeiten der Rechenregeln voneinander fragen, als sie für Ternärkörper ohne Topologie bestehen. Besonders starke derartige Verbindungen werden sich ergeben, wenn die Geraden der Ebene homöomorph zu einer reellen projektiven Geraden sind. Schließlich läßt sich auch noch die Frage nach der Einbettbarkeit topologischer Ternärkörper in topologische projektive Ebenen, d.h. nach der Existenz solcher Ebenen, die einen zu dem gegebenen topologischen Ternärkörper stetig isomorphen Ternärkörper haben, weiter fördern. Bisher kannte man die Antwort darauf in zwei Extremfällen, nämlich für die (algebraisch besonders „anständigen") Schiefkörper und für die (topologisch „anständigen") zusammenhängenden lokalkompakten Ternärkörper: Jeder topologische Schiefkörper läßt sich als Ternärkörper einer eindeutig bestimmten topologischen (desarguesschen)

projektiven Ebene auffassen (s. z. B. PICKERT [**1955**], S. 265). Jeder zusammenhängende lokalkompakte Ternärkörper gehört ebenso zu einer eindeutig bestimmten topologischen projektiven Ebene (SKORNJAKOV [**1954**]).

§ 7. Kompaktheit spezieller lokalkompakter Ebenen

Um in einer lokalkompakten projektiven Ebene die Kompaktheit der Geraden (und damit auch der Ebene) beweisen zu können, muß man eine Möglichkeit haben, genügend große kompakte Mengen angeben zu können, man muß also etwa die kompakten Umgebungen eines Punktes hinreichend stark „aufblähen" können. Das legt den folgenden Ansatz nahe: Man betrachtet für einen Punkt p, der nicht auf der Geraden G liegt, die Gruppe der (p, G)-Kollineationen. Kollineation soll eine eineindeutige Selbstabbildung der Ebene heißen, welche auch die Gesamtheit der Geraden (aufgefaßt als Punktmengen) umkehrbar eindeutig in sich überführt, (p, G)-Kollineation eine solche, die jede Gerade durch p und jeden Punkt auf G einzeln festläßt. Für die Gruppe der (p, G)-Kollineationen soll verlangt werden, daß sie auf den von p verschiedenen und nicht auf G liegenden Punkten einer jeden Geraden durch p transitiv sei. Diese (p, G)-Transitivität der Ebene bedeutet genau die Gültigkeit des desarguesschen Schließungssatzes mit Zentrum p und Achse G (s. z. B. PICKERT [**1955**], S. 76). Wählt man $p = 0$ und $G = u \cup v$, so hat die (p, G)-Transitivität der Ebene die Assoziativität der Multiplikation eines zu $0, u, v$ gehörigen Ternärkörpers zur Folge (PICKERT [**1955**], S. 104). Es zeigt sich, daß man für den Beweis der Kompaktheit von $0 \cup v$ mit einer viel schwächeren Voraussetzung auskommt: Die multiplikative Loop eines Ternärkörpers soll eine Alternativloop sein, d. h. je zwei Elemente der Loop sollen in einer assoziativen Unterloop (einer Untergruppe der Loop) liegen. Dann kann man sich wieder der Abbildung $x \to x' = x \backslash 1$ bedienen, um den um das Element $\infty = v$ vermehrten Ternärkörper mit zwei kompakten Mengen zu überdecken. Man geht dabei aus von einer kompakten Umgebung W von 0. Diese liefert bei dem Homöomorphismus $x \to x'$ eine kompakte Umgebung W' von ∞. Um jetzt die kompakte Menge W durch Multiplikation mit irgendeinem Element so zu strecken, daß sie das Komplement von W' umfaßt, sucht man erst einmal ein Element a mit der Eigenschaft, daß W echt in aW enthalten ist. Das besagt aber nichts anderes, als daß $a^{-1}W$ echt in W enthalten sein soll; denn aus $w = av$ folgt, daß w der von a und v erzeugten Loop angehört, also mit a und v zusammen in einer Gruppe liegt. Daher ist $w = av$ gleichwertig mit $a^{-1}w = v$ und $W \subset aW$ gleichwertig mit $a^{-1}W \subset W$. Ein a, das diese Forderung erfüllt, kann man aber leicht finden, da es ja, wie schon in § 3 bewiesen, zu jeder beliebig kleinen Umgebung U von 0 eine Umgebung V von 0 mit $VW \subseteq U$ gibt. Man braucht dann für $U \subset W$ nur $a = v^{-1}$ für ein von 0 verschiedenes Element v von V zu setzen. Jetzt ergibt sich durch eine Induktion, daß die Mengen $a^n W$ eine aufsteigende Folge von kompakten Umgebungen von 0 bilden. Da nämlich $a(a^n W)$ die Gesamtheit der Elemente $a(a^n w) = a^{n+1} w$ mit w in W ist, gilt $a(a^n W) = a^{n+1} W$ und aus $W \subset aW$ folgt durch wiederholte Multiplikation mit a von links $a^n W \subset a^{n+1} W$

für jedes n. Als homöomorphes Bild von W sind alle Mengen $a^n W$ kompakt. In der Folge dieser Mengen gibt es eine, die das Komplement von W' umfaßt; denn wäre das nicht der Fall, so könnte man für jedes n ein Element b_n finden, welches weder in $a^n W$ noch in W' läge. Für die zweite Bedingung kann man auch $b_n^{-1} \notin W$ schreiben, sie hätte also zur Folge, daß sich die b_n^{-1} nicht gegen 0 häufen könnten. Unter Berücksichtigung der Gruppeneigenschaft der von a und einem b_n erzeugten Loop bedeutet $b_n \notin a^n W$ dasselbe wie $a^{-n} b_n \notin W$. Nun hat aber die über a gemachte Voraussetzung die Konvergenz von a^{-n} gegen 0 zur Folge. Es ist nämlich $a^{-n} W \subset W$ für alle n, insbesondere liegen also alle a^{-n} in der kompakten Menge Ww^{-1}, wenn man für w irgendein von 0 verschiedenes Element aus W nimmt. Daher müssen die a^{-n} einen Häufungspunkt haben. Die Annahme, es gäbe einen solchen Häufungspunkt $c \neq 0$ führt auf einen Widerspruch: Gäbe es eine Teilfolge $\{n_i\}_{i=1}^{\infty}$ der n, für welche $\lim_{i \to \infty} a^{-n_i} = c \neq 0$ wäre, so wäre $\lim_{i \to \infty} a^{-n_{2i}} a^{n_i} = c c^{-1} = 1$ und die a^{-n} häuften sich auch gegen 1. Dann häuften sich aber die $a^{-n} w$ für jedes w aus W in w und jedes solche w läge als Häufungspunkt von Punkten aus der abgeschlossenen Menge $a^{-1} W$ selbst in $a^{-1} W$ im Widerspruch dazu, daß $a^{-1} W$ eine echte Teilmenge von W sein sollte. Jetzt kann man aus der ersten Bedingung über die b_n folgern, daß doch $\lim_{n \to \infty} b_n^{-1} = 0$ gelten müßte und damit ist dann gezeigt, daß es solche b_n überhaupt nicht geben kann. Für jedes feste n ist nämlich $\lim_{p \to \infty} a^{-p} b_n = 0$; für hinreichend große p liegt also $a^{-p} b_n$ in W. Es sei p_n die kleinste Zahl, für die das der Fall ist. Nach Konstruktion ist $p_n > n$, und $a^{-p_n} b_n$ kann nicht zu $a^{-1} W$ gehören, weil sonst p_n nicht minimal wäre. Daher liegen die $a^{-p_n} b_n$ alle in der Menge $W - a^{-1} W$, deren abgeschlossene Hülle K kompakt ist und den Punkt 0 nicht enthält ($a^{-1} W$ ist ja eine Umgebung von 0). Bei dem Übergang von x zu $x' = x^{-1}$ geht also K in die kompakte Teilmenge K' des Ternärkörpers T über. Es ist $b_n^{-1} \in K' a^{-p_n}$ und diese Mengen liegen schließlich in beliebig kleinen Umgebungen von 0, da K' kompakt und $\lim_{n \to \infty} a^{-p_n} = 0$ ist, also die $K' a^{-p_n}$ nach dem Beweis in § 3 einen gegen 0 konvergenten Filter bilden. Damit hat man bewiesen:

Ist die multiplikative Loop eines Ternärkörpers einer lokalkompakten projektiven Ebene eine Alternativloop oder ist insbesondere die Ebene (p, G)-transitiv für ein nicht inzidentes Paar p, G, so ist die Ebene kompakt.

§ 8. Topologisches Lemma

Für das Folgende wird ein topologisches Lemma gebraucht. Es bezeichne R den Raum der reellen Zahlen, S den Raum der reellen Zahlen modulo 1, also die eindimensionale Sphäre, n eine natürliche Zahl und q eine natürliche Zahl oder ∞; $T = R^n \times S^q$ sei das topologische Produkt von n Faktoren R und q Faktoren S.

Dann ist die Kompaktifizierung $T + \infty$ von T durch einen Punkt ∞ nicht homogen.

Beweis. T ist lokalkompakt, aber wegen $n \geq 1$ nicht kompakt. Daher läßt es sich auf genau eine Weise durch einen Punkt ∞ kompakt abschließen, und zwar so, daß man die Komplemente der kompakten Teilmengen von T in $T + \infty$ als Basis für den Umgebungsfilter in ∞ nimmt (BOURBAKI [1951], Top. gén. I, § 10, Th. 3). Jede kompakte Menge in R^n ist für ein geeignetes positives k in einer n-dimensionalen Kugel $K_n(k)$ vom Radius k enthalten, also jede kompakte Menge in T in einer geeigneten Menge $K_n(k) \times S^q$. Das Komplement von $K_n(k)$ in R^n hat die Form $(k, \infty) \times S_{n-1}$, wobei (k, ∞) das offene Intervall aller reellen Zahlen $> k$ bedeutet und S_{n-1} die $(n-1)$-dimensionale Sphäre. Die Mengen

$$U_k = (k, \infty) \times S_{n-1} \times S^q + \infty \quad (k \geq 1)$$

bilden also eine Basis für den Umgebungsfilter von ∞ bei der kompakten Abschließung von T durch ∞. Um die Inhomogenität nachzuweisen, vergleicht man die Umgebungsfilter in $0 = 0^n \times 0^q$ und ∞. Die Gesamtheit der Mengen

$$V = V(t, p) = (-t, t)^{n+p} \times S^r \quad (0 < t < \tfrac{1}{2},\ p \geq 1,\ p + r = q)$$

bildet eine Basis für den Umgebungsfilter in 0. Die Homogenität von $T + \infty$ würde insbesondere die Existenz eines Homöomorphismus σ bedeuten, der 0 in ∞ überführt. Dann muß es eine geeignete Menge V mit $V^\sigma \subseteq U_1$ geben und eine Zahl k mit $U_k \subseteq V^\sigma$. Wegen $0^\sigma = \infty$ gilt dann auch noch zwischen den punktierten Mengen die Einschachtelung

$$U_k - \infty \subseteq (V - 0)^\sigma = V^\sigma - \infty \subseteq U_1 - \infty.$$

Hieraus ergibt sich bei Betrachtung der Homotopieverhältnisse ein Widerspruch: Es sei $c > k$. Dann ist die Abbildung

$$x \to x^\gamma = c \times x \times 0^q$$

eine topologische Abbildung der $(n-1)$-Sphäre S_{n-1} in $U_k - \infty$. Das Bild S^γ_{n-1} ist in $U_1 - \infty$ nicht zusammenziehbar; denn eine Abbildung in einen Produktraum liefert genau dann ein zusammenziehbares Bild, wenn die Projektionen auf die Komponenten des Produktraumes zusammenziehbar sind. Die Projektion der Abbildung γ auf die Komponente S_{n-1} von $U_1 - \infty$ ist aber die identische Abbildung und eine Sphäre ist in sich nicht zusammenziehbar (KURATOWSKI [1952], S. 277; ALEXANDROFF-HOPF [1935], S. 513). Dagegen kann man S^γ_{n-1} in der vollen Umgebung U_k auf den Punkt ∞ zusammenziehen. Sei nämlich $x^{\gamma(s)} = c\,s \times x \times 0^q$ $(s \geq 1)$. Dann hängt $x^{\gamma(s)}$ stetig von x und s ab, und für $s \to \infty$ liegt $S^{\gamma(s)}_{n-1}$ schließlich in jeder Umgebung U_m von ∞. Daher ist S^γ_{n-1} erst recht in der größeren Menge V^σ zusammenziehbar. Nun hat aber die zu $R^{n+p} \times S^r$ homöomorphe Menge V^σ die Eigenschaft, daß jede in ihr gelegene zusammenziehbare $(n-1)$-Sphäre zusammenziehbar bleibt, wenn man aus V^σ einen nicht auf ihr gelegenen Punkt

herausnimmt. Sei nämlich φ die topologische Abbildung von S_{n-1} in $R^{n+p} \times S^r$ und $0^{n+p} \times 0^r = 0 \times 0$ der herauszunehmende Punkt; χ sei die Projektion von φ auf den Faktor R^{n+p}, ψ die Projektion auf S^r. Da 0×0 nicht auf S^φ_{n-1} liegen soll, ist $x^\chi \neq 0$ oder $x^\psi \neq 0$ für jeden Punkt x von S_{n-1}. Wegen der Umkehrbarkeit von φ ist außerdem $x^\chi \neq y^\chi$ oder $x^\psi \neq y^\psi$ für je zwei verschiedene Punkte x, y von S_{n-1}. Daher wird die Teilmenge N aller Punkte x von S_{n-1} mit $x^\psi = 0$ durch χ umkehrbar eindeutig und stetig in $R^{n+p} - 0$ abgebildet. N^χ ist eine höchstens $(n-1)$-dimensionale kompakte Teilmenge des punktierten mindestens $(n+1)$-dimensionalen euklidischen Raumes $R^{n+p} - 0$, läßt sich also dort auf einen Punkt zusammenziehen (ALEXANDROFF-HOPF [**1935**], S. 399, 405). Mit I als Bezeichnung für das abgeschlossene Intervall der reellen Zahlen zwischen 0 und 1 wird diese Zusammenziehung beschrieben durch eine stetige Abbildung α von $N \times I$ in $R^{n+p} - 0$, die auf $N \times 0$ in natürlicher Weise mit χ übereinstimmt und auf $N \times 1$ den Wert $a \neq 0$ annimmt. α läßt sich zu einer stetigen Abbildung β von $S_{n-1} \times I$ in R^{n+p} fortsetzen, die noch auf $S_{n-1} \times 0$ die Abbildung χ darstellt. Zieht man nun weiter die kompakte Menge $(S_{n-1} \times 1)^\beta$ im Raum R^{n+p} unter Festhalten des Punktes a stetig auf a zusammen, so hat man insgesamt die Menge S^χ_{n-1} in R^{n+p} stetig in den Punkt a übergeführt, ohne daß dabei ein Punkt aus N^χ durch 0 hindurchgegangen wäre. Anschließend kann man wegen der vorausgesetzten Zusammenziehbarkeit der Sphäre S^φ_{n-1} ihre Projektion S^ψ_{n-1} im Raum S^r stetig in einen Punkt b überführen. Diese beiden Deformationen hintereinander ergeben eine Zusammenziehung von S^φ_{n-1} auf den Punkt $a \times b$, die ganz in dem punktierten Raum $R^{n+p} \times S^r - 0 \times 0$ vor sich geht.

Damit ist die Bemerkung über die Zusammenziehbarkeit der $(n-1)$-Sphären in V^σ bewiesen, d. h. S^ψ_{n-1} ist zusammenziehbar in $V^\sigma - \infty$ und hieraus würde erst recht die Zusammenziehbarkeit in der umfassenderen Menge $U_1 - \infty$ folgen. Dieser Widerspruch beweist die Inhomogenität von $T + \infty$.

Bemerkung. In dem Sonderfall $n = 1$ läßt sich dieser Beweis viel einfacher führen. Dann besteht nämlich $S_{n-1} = S_0$ aus einem Punktepaar. Die Mengen $U_k - \infty$ haben daher genau zwei zusammenhängende Komponenten und jede Komponente von $U_1 - \infty$ enthält eine Komponente von $U_k - \infty$. Andererseits ist $(V-0)^\sigma$ zusammenhängend, also ganz in einer Komponente von $U_1 - \infty$ enthalten. Diese beiden Tatsachen zusammen widersprechen der Einschachtelung $U_k - \infty \subseteq (V-0)^\sigma \subseteq U_1 - \infty$.

§ 9. Die topologische Dimension lokalkompakter projektiver Ebenen

Das topologische Lemma soll nun ausgenutzt werden, um Aussagen über die topologische Dimension gewisser lokalkompakter projektiver Ebenen und über die Struktur ihrer additiven Loop zu machen. Für einen regulären Raum mit abzählbarer Basis (und ein solcher liegt ja bei einer lokalkompakten Ebene vor) wird die topologische Dimension rekursiv erklärt. Genau die leere Menge soll (-1)-dimensional sein, und ein Raum soll die Dimension n

haben, wenn jeder seiner Punkte eine Basis von Umgebungen mit $(n-1)$-dimensionalem Rand hat, aber nicht jeder Punkt eine solche mit $(n-2)$-dimensionalem Rand. Die anderen geläufigen Definitionen für eine topologische Dimension stimmen bei der betrachteten Raumklasse mit der hier gegebenen überein (HUREWICZ-WALLMANN [**1948**]). Für die nicht zusammenhängenden lokalkompakten Ebenen war in § 5 gezeigt worden, daß sie die Dimension 0 haben. Betrachtung der bekannten zusammenhängenden lokalkompakten Ebenen (etwa mit dem reellen oder komplexen Zahlkörper als Ternärkörper) läßt die Vermutung aufkommen, daß die Dimension dieser Ebenen eine endliche gerade Zahl sein muß. Da es beispielsweise einen eindimensionalen topologischen Raum gibt, dessen Produkt mit sich selbst wieder die Dimension 1 hat (ERDÖS [**1940**]), ist es auch bei endlicher Dimension einer topologischen Ebene nicht selbstverständlich, daß ihre Dimension gerade ist. Für beliebige lokalkompakte Ebenen ist die Vermutung noch nicht bestätigt, wohl aber gelingt das unter der Zusatzvoraussetzung, daß die Ebene einen Ternärkörper hat, dessen additive Loop eine kommutative Gruppe ist. Es handelt sich dann also um eine lokalkompakte kommutative Gruppe, die nach dem in § 5 und § 3 Bewiesenen zusammenhängend und lokalzusammenhängend ist und eine abzählbare Basis besitzt. Die Struktur dieser Gruppen ist vollständig bekannt (PONTRJAGIN [**1946**]). Mit R als der additiven Gruppe der reellen Zahlen und S als der additiven Gruppe der reellen Zahlen modulo 1 ist die additive Gruppe des Ternärkörpers das direkte Produkt von endlich vielen Faktoren R und höchstens abzählbar vielen Faktoren S, d.h. $T = R^n \times S^q$. Hierbei ist zunächst $0 \leq q \leq \omega_0$ und n eine natürliche Zahl; denn wäre $n = 0$, so entfiele der erste Bestandteil des direkten Produktes und T wäre als das Produkt der kompakten Mengen S selbst kompakt. Das kann aber nicht sein, da sonst der Punkt ∞, der T zur projektiven Geraden ergänzt, als Komplement einer kompakten und daher abgeschlossenen Menge offen wäre und somit die projektive Gerade in einem und daher in allen Punkten diskret, was von vornherein ausgeschlossen war. Nun ist aber die projektive Gerade $T + \infty$ ein homogener topologischer Raum, also kann nach dem topologischen Lemma nicht auch $0 < q$ sein, d.h. es folgt $q = 0$. Damit ist bewiesen:

Ist die additive Loop eines Ternärkörpers einer lokalkompakten zusammenhängenden projektiven Ebene eine kommutative Gruppe, so ist sie für eine geeignete natürliche Zahl n das n-fache direkte Produkt der additiven Gruppe der reellen Zahlen. Die projektive Ebene hat dann die topologische Dimension $2n$.

Insbesondere gibt es also dann in der additiven Gruppe außer der 0 keine Elemente endlicher Ordnung und das ergibt eine Folgerung über Antifanoebenen, das sind solche Ebenen, in denen jedes vollständige Viereck kollineare Diagonalpunkte hat. In einer solchen Ebene erfüllt jedes Element x der additiven Loop die Gleichung $x + x = 0$ und hieraus folgt das kommutative und assoziative Gesetz für die Addition (PICKERT [**1955**], S. 60). Also:

Eine lokalkompakte Antifanoebene ist notwendig nulldimensional.

§ 10. Projektive Ebenen, die zur reellen projektiven Ebene homöomorph sind

Die erste Frage, die sich bei zur reellen Ebene homöomorphen projektiven Ebene erhebt, ist die, ob ihre Geraden auch zur reellen Geraden homöomorph sind. Diese Frage ist zu bejahen. Beweis. Zunächst weiß man aus § 5, daß die projektiven Geraden einer zur reellen Ebene homöomorphen Ebene zusammenhängende und lokalzusammenhängende Mengen sind, denen als abgeschlossenen Teilmengen eines kompakten Raumes mit abzählbarer Basis dieselben Eigenschaften zukommen. Von einem solchen Raum weiß man (KURATOWSKI [**1952**], S. 184), daß er sogar lokal kurvenmäßig zusammenhängend ist, d.h. es gibt (beliebig kleine) Umgebungen eines Punktes x, so daß jeder Punkt y einer solchen Umgebung mit x durch eine zu dem reellen Intervall $I=[0,1]$ homöomorphe Menge (d.i. eine einfache Kurve) in dieser Umgebung verbunden ist. Die doppelte Transitivität der Homöomorphismengruppe der affinen Geraden ergibt daraus sofort, daß in der affinen Geraden je zwei Punkte durch eine einfache Kurve verbunden sind. Es wird sich jetzt sogar zeigen, daß diese einfache Kurve eindeutig bestimmt ist. Gäbe es nämlich zwei verschiedene einfache Kurven, die zwei Punkte miteinander verbinden, so erhielte man nach Weglassen geeigneter Kurvenstücke eine einfache geschlossene Kurve, d.h. eine Jordan-Kurve in der affinen Geraden. Diese Jordan-Kurve läßt sich nun innerhalb der affinen Geraden auf einen Punkt zusammenziehen, und zwar durch eine Schar von (bis auf den Schluß) homöomorphen Abbildungen. Sei nämlich die Jordan-Kurve dargestellt durch $a(s)$ $(s\in S)$ und eine die Punkte 0 und 1 verbindende einfache Kurve durch $b(t)$ $(t\in I)$ mit $b(0)=1$ und $b(1)=0$. Dann ist $c_t(s)=a(s)b(t)$ stetig in s, t und $c_0(s)=a(s)$, $c_1(s)=0$. Da nun die betrachtete Ebene lokaleuklidisch ist, liegt eine hinreichend kleine Jordan-Kurve der affinen Geraden ganz in einer offenen euklidischen Kreisscheibe, teilt diese also in ein Inneres und ein Äußeres. Wegen der Zusammenziehbarkeit der Jordan-Kurve in der affinen Geraden muß das Innere der Jordan-Kurve ganz der affinen Geraden angehören, dann wäre aber die Gerade zweidimensional und die Ebene vierdimensional, ein Widerspruch.

Spricht man dieses Zwischenergebnis, daß in der affinen Geraden je zwei Punkte eindeutig durch eine einfache Kurve verbunden sind, projektiv aus, so erhält man die Aussage: Zu drei (paarweise verschiedenen) Punkten p, q, r gibt es in der projektiven Geraden genau eine einfache Kurve K mit den Endpunkten p, q, welche r nicht enthält. Ist 0 irgendein von p und q verschiedener Punkt aus K, so gibt es auch genau eine einfache Kurve L mit den Endpunkten p, q, die 0 nicht enthält. Diese muß dann durch r und ebenso natürlich durch jeden anderen nicht auf K gelegenen Punkt hindurchgehen. Die beiden einfachen Kurven K und L füllen also zusammen die ganze projektive Gerade aus. Sie haben aber nur die Endpunkte p, q gemeinsam; denn gäbe es einen von p, q verschiedenen Punkt x, der zugleich K und L angehörte, so müßte die einfache Kurve M, die p mit q verbindet und x nicht enthält, sowohl durch 0 wie durch r hindurchgehen, da sie ja von K und L

verschieden ist. Nun gibt es eine Projektivität, die 0 mit p und r mit q vertauscht (man kann sie leicht durch Hintereinander-Ausführung von drei Projektionen angeben). Diese würde aus M eine einfache Kurve mit den Endpunkten $0, r$ machen, die p und q enthielte. Weglassen der Endstücke ergäbe eine einfache Kurve von p nach q, die weder durch 0 noch durch r hindurchginge, also wieder von K und L verschieden wäre. Eine solche kann es aber nicht geben. Damit ist gezeigt, daß die projektive Gerade eine einfache geschlossene Kurve ist, also:

Die projektiven Geraden einer zur reellen Ebene homöomorphen projektiven Ebene sind eindimensionale Sphären, die affinen Geraden sind homöomorph zur Menge der reellen Zahlen.

§ 11. Topologische Loops, die zur reellen Geraden homöomorph sind

Ist ein Ternärkörper homöomorph zur reellen Geraden, so bestehen zwischen den in ihm möglichen Rechenregeln engere Abhängigkeiten, als bei anderen lokalkompakten Ternärkörpern. Das wird dadurch bedingt, daß die von den reellen Zahlen übertragene Anordnung in gewisser Weise mit den Ternärkörper-Verknüpfungen verträglich ist. Zunächst werde nur die additive Loop betrachtet. Von ihr soll vorausgesetzt werden, daß jedes ihrer Elemente eine assoziative Untergruppe der Loop erzeugt; eine solche Loop nennt man auch potenzassoziativ (im Gedanken an multiplikativ geschriebene Loops). In Anlehnung an BOURBAKI ([**1947**], Top. gén. V, § 2) ergibt sich, daß eine solche zur Menge der reellen Zahlen homöomorphe topologische Loop angeordnet und sogar archimedisch angeordnet ist. Daß es sich um eine angeordnete Loop handelt, bedeutet, daß aus $x<y$ folgt $x+z<y+z$ und $z+x<z+y$. Dabei werde die Beziehung $x>0$ durch den Homöomorphismus von den reellen Zahlen her übertragen und $x<y$ definiert durch $y-x>0$, wobei $y-x$ dasjenige Element der Loop sein soll, das $(y-x)+x=y$ erfüllt. Dann ist die durch $f(z)=(y+z)-(x+z)$ gegebene Abbildung f, da die Loopoperationen ja stetig sind, eine stetige Funktion. Im Fall $x<y$ ist $f(0)>0$ und $f(z)\neq 0$ für jedes z, da ja $f(z)=0$ dasselbe bedeutet wie $y+z=x+z$ oder $y=x$. Nun bildet f die (zusammenhängende) Loop auf eine zusammenhängende Punktmenge ab, die ein Element >0 enthält, aber nicht die 0 selbst, also nur aus Elementen >0 besteht. Damit ist die Rechtsmonotonie der Addition bewiesen. Jetzt folgen nachträglich auch die Anordnungsgesetze für die Beziehung $<$. Die Linksmonotonie ergibt sich ganz analog. Für eine natürliche Zahl n wird $a\cdot n$ induktiv definiert durch $a\cdot 1=a$, $a\cdot(n+1)=a\cdot n+a$. Ist $0<a$ und $0<b$ und wäre $a\cdot n\leq b$ für jedes n, so konvergierte die monoton wachsende Folge $\{a\cdot n\}_{n=1}^{\infty}$ gegen ein Element $c\leq b$ und dann müßte die Folge $\{a\cdot(n+1)\}_{n=1}^{\infty}$ sowohl gegen c wie auch gegen $c+a$ konvergieren. Also gibt es Zahlen n mit $a\cdot n>b$. Entsprechendes gilt für $b<a<0$, d. h. die Loop ist archimedisch angeordnet.

Jetzt läßt sich in Verschärfung des Satzes von PICKERT [**1955**], S. 246 beweisen:

Eine archimedisch angeordnete potenzassoziative Loop ist isomorph zu einer Untergruppe der additiven Gruppe der reellen Zahlen.

Der Beweis entspricht dem bei BOURBAKI [**1947**], Top. gén. V, § 2 für Gruppen gegebenen. Es sind zwei Fälle zu unterscheiden. Im ersten Fall gibt es ein Element e, das unmittelbarer Nachfolger von 0 ist. Dann folgt durch eine naheliegende Induktion, daß $e \cdot (n+1)$ unmittelbarer Nachfolger von $e \cdot n$ ist für jedes ganze n und wegen der archimedischen Anordnung erschöpfen die Elemente $e \cdot n$ die ganze Loop, die also isomorph zur Gruppe der ganzen Zahlen ist. Im zweiten Fall gibt es keinen unmittelbaren Nachfolger von 0 und die Loop ist dicht geordnet. Für Elemente $x>0$ und beliebige y werde $y:x$ als die größte (positive und negative) ganze Zahl n mit $x \cdot n \leq y$ erklärt. Wählt man noch ein festes Element $e>0$, so ist

$$a^{\varrho} = \lim_{x \to 0} \frac{a:x}{e:x}$$

für jedes Element a vorhanden und >0 genau für $a>0$, wie eine leichte, für $a>0$ und $a<0$ getrennt durchzuführende Rechnung zeigt, die genau mit der bei BOURBAKI angegebenen übereinstimmt. Wegen $a:x+b:x \leq (a+b):x \leq a:x+b:x+2$ ist ϱ ein Homomorphismus der Loop in die additive Gruppe der reellen Zahlen, und da aus $a<b$ die Existenz eines Elementes $c>0$ mit $a+c=b$, also $a^{\varrho}+c^{\varrho}=b^{\varrho}$ und $a^{\varrho}<b^{\varrho}$ folgt, ist ϱ sogar ein Isomorphismus. Weiter gilt:

Eine zur Gruppe der reellen Zahlen homöomorphe potenzassoziative Loop ist sogar mit ihr isomorph;

denn die Gruppe der reellen Zahlen hat keine zusammenhängende echte Untergruppe.

§ 12. Zur reellen Ebene homöomorphe Translationsebenen

Eine Ebene, in der nach Auszeichnung einer uneigentlichen Geraden die Translationen transitiv sind auf der Menge der eigentlichen Punkte, heißt Translationsebene. Ihr Ternärkörper (T, τ) (mit u, v auf der uneigentlichen Geraden) ist ein Quasikörper, d.h. außer den Auflösbarkeitsbedingungen gilt $\tau(s, x, t) = sx+t$, mit der Addition bildet T eine Gruppe und es ist stets $a(b+c) = ab+ac$. Nun gilt:

Eine zur reellen Ebene homöomorphe Translationsebene ist mit der reellen Ebene isomorph.

Beweis. Die additive Gruppe des Quasikörpers stimmt mit der additiven Gruppe des reellen Körpers überein. Kern des Quasikörpers nennt man die Menge aller Elemente x, die $(ab)x = a(bx)$ und $(a+b)x = ax+bx$ für alle Paare a, b erfüllen. Der Kern eines Quasikörpers ist ein Schiefkörper, der das Element 1 enthält. Wegen der Stetigkeit der Abbildungen $x \to (ab)x - a(bx)$ und $x \to (a+b)x - (ax+bx)$ ist der Kern eines topologischen Quasikörpers als Durchschnitt von Urbildern der abgeschlossenen Menge $\{0\}$ abgeschlossen. Mit 1 liegt nun der ganze von 1 erzeugte Körper der rationalen Zahlen (die

Charakteristik ist 0 wegen der Anordnung) im Kern. Diese Menge liegt aber überall dicht in dem Quasikörper, der daher mit seinem Kern übereinstimmt und der Körper der reellen Zahlen ist.

§ 13. Zur reellen Ebene homöomorphe Neokörperebenen

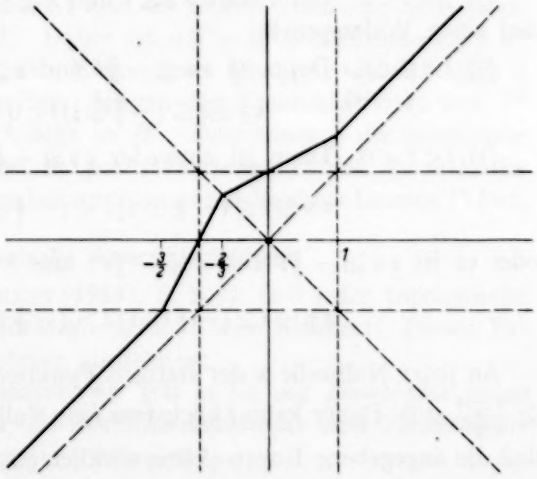
Nach dem Ergebnis der beiden letzten Paragraphen genügt [neben der „Zerlegbarkeitsbedingung" $\tau(s, x, t) = sx + t$] schon die Forderung der Potenzassoziativität der Addition und das eine Distributivgesetz, um für einen zum Körper der reellen Zahlen homöomorphen Ternärkörper zu erzwingen, daß er mit diesem isomorph ist. Auf die Forderung der Potenzassoziativität der Addition kann man dabei auch dann nicht verzichten, wenn man sehr viel stärkere Bedingungen für die Multiplikation voraussetzt, wie das folgende Beispiel zeigt:

Es gibt einen zur Menge der reellen Zahlen homöomorphen echten kommutativen (planaren) Neokörper mit kommutativer und kürzbarer Addition.

Als Neokörper wird dabei ein Ternärkörper (T, τ) bezeichnet mit $\tau(s, x, t) = sx + t$ und assoziativer und beiderseits distributiver Multiplikation. Ein kommutativer Neokörper erfüllt außerdem noch das kommutative Gesetz der Multiplikation. Die Kürzbarkeit bedeutet bei kommutativer Addition die Gültigkeit von $(-x) + (x + y) = y$. Mit dem Beispiel ist gleichzeitig die Frage von PICKERT [**1956**] nach der Existenz echter Neokörper mit Kürzungsregel beantwortet. Das Beispiel ergibt sich durch geeignete Anwendung des Gedankens von MOULTON [**1902**], die Geraden der gewöhnlichen reellen Ebene durch passende geknickte Linien zu ersetzen. Man erhält es dadurch, daß man die Multiplikation des reellen Körpers beibehält und eine neue Addition $\oplus$ einführt durch die Definition

$$\begin{aligned} x \oplus t &= x + \tfrac{1}{2}t \quad \text{für} \quad t^{-1}x \leq -\tfrac{3}{2} \quad \text{oder} \quad t^{-1}x \geq 1 \\ &= \tfrac{1}{2}x + t \quad \text{für} \quad -\tfrac{2}{3} \leq t^{-1}x \leq 1 \\ &= 2x + 2t \quad \text{für} \quad -\tfrac{3}{2} \leq t^{-1}x \leq -\tfrac{2}{3}. \end{aligned}$$

Man sieht sofort, daß die neue Addition kommutativ ist und rechnet leicht $(-1) \oplus (1 \oplus x) = x$ nach, woraus mit Hilfe der Distributivität, die sich unmittelbar aus der Definition ergibt (die Fallunterscheidung richtet sich nur nach dem Wert von $t^{-1}x$), die Kürzungsregel folgt. In der Figur bedeutet die getroffene Festsetzung, daß die geknickte „Gerade" die Gleichung $y = x \oplus 1$ erhält. Alle anderen Geraden außer den



Achsenparallelen erhält man daraus durch Streckung in der x- und y-Richtung.

Daß man so tatsächlich einen Neokörper erhält, erfordert den Nachweis (PAIGE [1950]), daß für $s \neq s'$ die Gleichung $sx \oplus t = s'x \oplus t'$ eindeutig nach x auflösbar ist. Hierzu betrachtet man die stetige reelle Funktion f mit $f(x) = (sx \oplus t) - (s'x \oplus t')$. Ist $s = 0$ oder $s' = 0$ oder haben s und s' verschiedenes Vorzeichen, so ist f eine streng monotone Funktion, deren Ableitung, abgesehen von drei Stellen, überall erklärt ist und für die $|f'(x)| = |s - s'| > 0$ bei genügend großem Absolutwert von x gilt. Daher hat f in diesen Fällen genau eine Nullstelle. In den anderen Fällen kann man durch Wahl der Bezeichnung erreichen, daß s' der absolut kleinere der beiden Werte s, s' ist. Da die Auflösung der betrachteten Gleichung gleichwertig ist mit der Auflösung der aus ihr durch Multiplikation mit s'^{-1} hervorgehenden Gleichung, kann man sich auf $s' = 1 < s$ beschränken. Im Fall $t = t'$ ist nichts mehr zu beweisen; sonst kann man auch noch annehmen, daß $t < t'$ ist; der gegenteilige Fall läßt sich nämlich durch Wechsel des Vorzeichens von x und Multiplikation der Gleichung mit (-1) hierauf zurückführen. Für hinreichend großen Absolutwert von x gilt immer der erste Fall der Definition von $\oplus$ und es ist $f'(x) = s - 1 > 0$. Daher existiert mindestens eine Nullstelle von f. Mittels der Monotoniegesetze für die Verknüpfung $\oplus$ folgt wegen der vorausgesetzten Ungleichungen für s, t, t' aus $x \leqq 0$ stets $f(x) < 0$. Man kann sich also auf die Betrachtung der Funktion f im Intervall $x > 0$ beschränken. Für einen Wert x, bei dem nicht $f'(x) \geqq \frac{s-1}{2}$ gilt, bei dem also ein Knick vorliegt oder die Ableitung einen Wert $< \frac{s-1}{2}$ hat, liegt entweder für die Definition von $sx \oplus t$ der zweite Fall vor, aber für die Definition von $x \oplus t'$ der erste oder dritte Fall, oder $sx \oplus t$ ist nach dem ersten oder zweiten Fall definiert und $x \oplus t'$ nach dem dritten. Bei jeder dieser Möglichkeiten ergibt sich $f(x) \neq 0$, und zwar so: Sei

(1) $0 < t < t'$. Dann wäre $sx \leqq t$ und $x \geqq t'$. Dies führt aber wegen $x < sx$ auf einen Widerspruch.

(2) $t < 0 < t'$. Dann ist $sx \leqq -\frac{2}{3}t$ und $x \geqq t'$, also

$$sx \oplus t \leqq (-\tfrac{2}{3} \oplus 1)\, t < 0 < x \oplus t'.$$

(3) $t < t' < 0$. Dann ist entweder $sx \leqq -\frac{2}{3}t$ und $x \geqq -\frac{2}{3}t'$, also

$$sx \oplus t \leqq (-\tfrac{2}{3} \oplus 1)\, t < (-\tfrac{2}{3} \oplus 1)\, t' \leqq x \oplus t';$$

oder es ist $sx \geqq -\frac{3}{2}t$ und $x \leqq -\frac{3}{2}t'$, also

$$sx \oplus t \geqq (-\tfrac{3}{2} \oplus 1)\, t > (-\tfrac{3}{2} \oplus 1)\, t' \geqq x \oplus t'.$$

An jeder Nullstelle x der stetigen Funktion f ist also $f'(x)$ vorhanden und $\geqq \frac{s-1}{2} > 0$. Daher kann f höchstens eine Nullstelle haben. Damit ist gezeigt, daß die angegebene Konstruktion wirklich einen (planaren) Neokörper liefert.

Kann man also im allgemeinen bei einem zur Menge der reellen Zahlen homöomorphen Neokörper nichts über die Addition aussagen, so gilt doch wenigstens für die Multiplikation:

Die multiplikative Gruppe $T^\times$ eines zum Körper der reellen Zahlen homöomorphen Neokörpers T ist isomorph zur multiplikativen Gruppe $R^\times$ dieses Körpers.

Da es nichtkommutative zu $R^\times$ homöomorphe Gruppen gibt (etwa die Erweiterung der Gruppe der positiven reellen Zahlen durch den Automorphismus $x \to x^{-1}$), aber wie sich zeigen wird, alle kommutativen zu $R^\times$ homöomorphen Gruppen mit $R^\times$ isomorph sind, handelt es sich beim Beweis im wesentlichen um den Nachweis der Kommutativität von $T^\times$. Zunächst ergibt sich aus den bekannten Sätzen über topologische Gruppen (BOURBAKI [**1947**], Top. gén. V, § 2), daß die zusammenhängende Komponente H des Elementes 1 von T isomorph ist zur Gruppe der positiven reellen Zahlen. Nun betrachtet man die durch $x + x^\varepsilon = 0$ erklärte Abbildung ε. Wegen $0^\varepsilon = 0$ stellt ε einen Homöomorphismus von $T^\times$ dar, vertauscht also die beiden zusammenhängenden Komponenten von $T^\times$ oder läßt sie fest. Das letzte kann aber nicht der Fall sein; denn die additive Loop $(T, +)$ ist nach § 11 eine angeordnete Loop und H besteht bei der dort eingeführten Anordnung aus allen Elementen > 0, also ist $H + H \subseteq H$, während $H + H^\varepsilon$ das Element 0 enthält, so daß also H und H^ε verschieden sind und $T^\times = H \cup H^\varepsilon$ ist. Wegen $(a + a^\varepsilon) b = b(a + a^\varepsilon) = 0$ gilt $ab + a^\varepsilon b = ba + ba^\varepsilon$. Für a, b in H folgt also aus der Kommutativität von H auch $a^\varepsilon b = ba^\varepsilon$ und mit Hilfe dieser Beziehung ergibt sich auf dieselbe Weise $a^\varepsilon b^\varepsilon = b^\varepsilon a^\varepsilon$, so daß also $T^\times$ kommutativ ist.

Um jetzt den Isomorphismus zu den reellen Zahlen von der zusammenhängenden Komponente H von 1 auf die ganze Gruppe $T^\times$ auszudehnen, sucht man ein Element e im Komplement H' von H in $T^\times$ mit $e^2 = 1$. Die Abbildung $x \to ax$ mit einem festen Element a aus H' ist ein Homöomorphismus von $T^\times$ und führt H in eine zusammenhängende Menge über, die das Element a enthält, sie vertauscht also die beiden zusammenhängenden Komponenten H und H' von T. Daher ist $aH' = H$ und insbesondere a^2 ein Element h aus H. In H ist jedes Element ein Quadrat, also läßt sich h in der Form b^2 mit b in H schreiben. Wegen der Kommutativität von $T^\times$ ist jetzt $(ab^{-1})^2 = 1$ und $e = ab^{-1}$ liegt in H'. Bezeichnet ϱ die isomorphe Abbildung von H auf die Gruppe der positiven reellen Zahlen, so setzt die Abbildung $ex \to -x^\varrho$ für x in H den Isomorphismus auf die ganze Gruppe $T^\times$ fort.

§ 14. Topologische Alternativkörper

Es ist bekannt (s. z.B. PICKERT [**1955**], S. 265), daß jeder topologische Schiefkörper eine topologische (desarguessche) Ebene definiert. Dieses Ergebnis läßt sich auf Alternativkörper ausdehnen:

Zu einem topologischen Alternativkörper gibt es bis auf Isomorphie genau eine topologische projektive Ebene, die den Alternativkörper zum Ternärkörper hat und in ihm die gegebene Topologie induziert.

Beweis. Nach dem üblichen Verfahren wird zu dem Alternativkörper T eine projektive Ebene konstruiert, indem man alle Paare (x, y) mit $x, y \in T$ als eigentliche Punkte nimmt, die Mengen $[s, t]$ der Punkte (x, y) mit $sx + t = y$ und die Mengen $[r]$ der Punkte (r, y) mit $y \in T$ als affine Geraden. Dazu kommen die Punkte (z), die eingeführt werden als Schnittpunkte aller Geraden $[z, t]$ mit $t \in T$ und der Punkt (∞) als Schnittpunkt der Geraden $[r]$ mit $r \in T$ als uneigentliche Punkte, die alle zusammen die uneigentliche Gerade $[\infty]$ bilden.

Wenn es überhaupt eine Topologisierung der so gebildeten projektiven Ebene gibt, die auf T die gegebene Topologie hervorruft, so ist diese Topologie der Ebene eindeutig durch die von T bestimmt, und zwar so, daß die affine Ebene topologisches Produkt von T mit sich selbst wird. Das bedeutet, daß die Gesamtheit der Mengen $U \times V$, wo U Umgebung von x und V Umgebung von y ist, eine Basis für den Umgebungsfilter des Punktes (x, y) bilden muß. Weiter werden jetzt die Umgebungen für die uneigentlichen Punkte und die Geraden in geeigneter Weise definiert und dann wird nachgerechnet, daß bei diesen Umgebungsdefinitionen Verbinden und Schneiden wirklich stetige Operationen werden. Damit sind dann die Umgebungsdefinitionen auch gerechtfertigt. Für den Umgebungsfilter des Punktes (z) nimmt man als Basis die Gesamtheit der Mengen, die für eine Umgebung U von 0 und eine Umgebung V von z aus allen Punkten (x, y) mit $x^{-1} \in U$ und $yx^{-1} \in V$ und allen Punkten (w) mit $w \in V$ bestehen. Für den Umgebungsfilter des Punktes (∞) nimmt man als Basis die Gesamtheit der Mengen, die für eine Umgebung U von 0 aus allen Punkten (x, y) mit $y^{-1} \in U$ und $xy^{-1} \in U$ und allen Punkten (w) mit $w^{-1} \in U$ bestehen. Die Umgebungen für die Geraden gehen daraus durch Vertauschen von x mit s und y mit $-t$ bei Umkehrung der Multiplikationsreihenfolge und Ersetzen der runden Klammern durch eckige hervor, womit eine Dualität zwischen Punkt- und Geradentopologie gegeben ist. Wenn man die Stetigkeit des Verbindens nachgerechnet hat, ergibt sich mit Hilfe dieser Dualität auch die des Schneidens.

Der Nachweis der Stetigkeit muß für die verschiedenen Typen von Verbindungsgeraden getrennt geführt werden, verläuft jedoch immer nach dem gleichen Schema.

(1) Verbindung von (x_1, y_1) und (x_2, y_2) mit $x_1 \neq x_2$.

Die Verbindungsgerade ist vom Typ $[s, t]$ und bestimmt sich durch $y_i = sx_i + t$ $(i = 1, 2)$. Es ist $s = (y_1 - y_2)(x_1 - x_2)^{-1}$, $t = y_1 - sx_1$. Also hängen s, t wegen $x_1 - x_2 \neq 0$ stetig von x_1, y_1, x_2, y_2 ab.

(2) Verbindung von (p, q_1) und (p, q_2).

Die Verbindungsgerade ist $[p]$. Eine Umgebung von $[p]$ ist bestimmt durch Angabe einer Umgebung U von 0 und einer Umgebung V von p und besteht aus allen Geraden $[s, t]$ mit $s^{-1} \in U$, $-s^{-1}t \in V$ und allen Geraden $[r]$ mit $r \in V$. Man muß Umgebungen W_i von (p, q_i) finden, so daß alle Verbindungsgeraden eines Punktes (x_1, y_1) von W_1 mit einem Punkt (x_2, y_2) von W_2 in

einer solchen vorgegebenen Umgebung von $[p]$ liegen. Ist $x_1 \neq x_2$ und hat man W_1, W_2 so klein gewählt, daß stets $y_1 \neq y_2$ ist, so gilt $s^{-1} = (x_1 - x_2)(y_1 - y_2)^{-1}$, $-s^{-1}t = x_1 - s^{-1}y_1$. Daher kann man durch geeignete Wahl von W_1 und W_2 erreichen, daß $s^{-1} \in U$ und $-s^{-1}t \in V$. Ist $x_1 = x_2 = r$, so ist $[r]$ die Verbindungsgerade der beiden Punkte, und indem man die W_i gegebenenfalls noch verkleinert, erreicht man auch noch $r \in V$.

(3) Verbindung von (p, q) und (z).

Die Verbindungsgerade ist $[z, t]$ mit $q = zp + t$. Zu einer Umgebung U von z und einer Umgebung V von $-zp + q$ muß man also eine Umgebung W von (p, q) und eine Umgebung Z von (z) angeben, so daß die Verbindungsgerade eines Punktes (x, y) aus W mit einem Punkt (u, v) aus Z oder einem Punkt (w) aus Z von der Form $[s, t]$ mit $s \in U$, $t \in V$ ist. Daß die Verbindungsgerade vom Typ $[s, t]$ ist, bedeutet $x \neq u$. Man kann sich W und Z von vornherein so klein gewählt denken, daß das stets der Fall ist. Aus $y = sx + t$ und $v = su + t$ folgt dann, wenn man t eliminiert, die Gleichung $s - (sx)u^{-1} = vu^{-1} - yu^{-1}$. Ihre rechte Seite soll mit a abgekürzt werden. Weiter erhält man hieraus

$$s = (au)(u - x)^{-1} = a + (ax)(u - x)^{-1} = a + (ax)\left(u^{-1}(1 - xu^{-1})^{-1}\right)$$

und, wenn man für a seinen Wert einsetzt,

$$s = vu^{-1} - yu^{-1} + \left((vu^{-1} - yu^{-1})x\right)\left(u^{-1}(1 - xu^{-1})^{-1}\right).$$

Die Umgebung Z von (z) ist dadurch gegeben, daß u^{-1} in einer Umgebung von 0 und vu^{-1} in einer Umgebung von z liegen, und Z wird dadurch beliebig klein, daß man diese Umgebungen beliebig klein wählt. Nun hängt s stetig von x, y, u^{-1}, vu^{-1} ab, und wenn nur W und Z genügend klein sind, so ist $s \in U$, und man kann auch noch erreichen, daß $t = -sx + y$ in V liegt.

Die Verbindungsgerade von (x, y) mit (w) hat die Form $[w, t]$, wo w in einer durch Z bestimmten Umgebung von z liegt und $y = wx + t$ ist. Für hinreichend kleines W und Z gilt also $w \in U$ und $t \in V$. Daher kann man W und Z so klein wählen, daß gleichzeitig diese und die vorher gestellten Forderungen erfüllt werden.

(4) Verbindung von (p, q) und (∞).

Die Verbindungsgerade ist $[p]$. Eine Umgebung von $[p]$ ist wie in (2) bestimmt. Man muß eine Umgebung W von (p, q) und eine Umgebung Z von (∞) finden, so daß die Verbindungsgeraden vom Typ $[s, t]$ eines Punktes (x, y) aus W mit einem Punkt (u, v) aus Z oder einem Punkt (w) aus Z die Bedingungen $s^{-1} \in U$, $-s^{-1}t \in V$ erfüllen und die Geraden vom Typ $[r]$ die Bedingung $r \in V$.

Ist $x \neq u$, so ist die zugehörige Verbindungsgerade vom Typ $[s, t]$. Aus $s = 0$ folgt $y = v$. Da aber die Umgebungen von (∞) dadurch gekennzeichnet sind, daß v^{-1} und uv^{-1} in einer Umgebung von 0 liegen, kann man W und Z von vornherein so klein gewählt denken, daß $y \neq v$, also $s \neq 0$ ist. Die Gleichungen $y = sx + t$ und $v = su + t$ kann man dann umformen zu $s^{-1}y = x + s^{-1}t$

und $s^{-1}=uv^{-1}+(s^{-1}t)v^{-1}$. Elimination von $s^{-1}t$ ergibt $s^{-1}-(s^{-1}y)v^{-1}=uv^{-1}-xv^{-1}$. Eine entsprechende Umformung wie bei (3) führt auf die Gleichung

$$s^{-1}=uv^{-1}-xv^{-1}+\big((uv^{-1}-xv^{-1})y\big)\big(v^{-1}(1-yv^{-1})^{-1}\big).$$

Wegen $y\neq v$ ist hier $1-yv^{-1}\neq 0$ und daher dieser Ausdruck erklärt. s^{-1} hängt also stetig von x, y, v^{-1}, uv^{-1} ab und für hinreichend kleines W und Z ist $s^{-1}\in U$ und $-s^{-1}t=-s^{-1}y+x\in V$.

Ist $x=u=r$, so ist $[r]$ die zugehörige Verbindungsgerade und für eine hinreichend kleine Umgebung W ist $r\in V$.

Die Verbindungsgerade von (x, y) mit (w) ist von der Form $[w, t]$ mit $y=wx+t$. Für eine hinreichend kleine Umgebung Z von (∞) liegt w^{-1} in der betrachteten Umgebung U von 0 und bei passender Wahl von W kann man auch noch $-w^{-1}t=-w^{-1}y+x\in V$ erreichen.

(5) Verbindung von (z_1) und (z_2).

Die Verbindungsgerade ist $[\infty]$. Eine Umgebung von $[\infty]$ besteht aus allen Geraden $[s, t]$ mit $t^{-1}\in U$, $t^{-1}s\in U$ und allen Geraden $[r]$ mit $r^{-1}\in U$, wo U eine Umgebung von 0 ist. Man muß Umgebungen Z_i von (z_i) finden, so daß die Verbindungsgeraden eines Punktes aus Z_1 mit einem Punkt aus Z_2 stets in der gegebenen Umgebung von $[\infty]$ liegen. Z_i besteht aus den Punkten (u_i, v_i), wo u_i^{-1} in einer Umgebung von 0 und $v_iu_i^{-1}$ in einer Umgebung von z_i liegt, und aus den Punkten (w_i) mit w_i aus einer Umgebung von z_i. Die Verbindungsgerade von (u_1, v_1) mit (u_2, v_2) ist im Fall $u_1=u_2=r$ von der Form $[r]$ und für hinreichend kleine Umgebungen Z_i liegt r^{-1} in U. Im Fall $u_1\neq u_2$ erhält man aus $v_i=su_i+t$ $(i=1, 2)$ durch Elimination von s die Gleichung $t^{-1}=(u_1^{-1}-u_2^{-1})(v_1u_1^{-1}-v_2u_2^{-1})^{-1}$ und weiter $t^{-1}s=t^{-1}(v_1u_1^{-1})-u_1^{-1}$. Für hinreichend kleine Umgebungen Z_i liegen t^{-1} und $t^{-1}s$ in U. Die Verbindungsgerade von (u, v) aus Z_1 mit (w) aus Z_2 ist $[w, t]$ mit $v=wu+t$. Hieraus berechnet sich $t^{-1}=u^{-1}(vu^{-1}-w)^{-1}$ und $t^{-1}w=t^{-1}(vu^{-1})-u^{-1}$. Für hinreichend kleine Z_i liegt also wieder die Verbindungsgerade in der gegebenen Umgebung von $[\infty]$.

Die Verbindungsgerade von $(w_1)\in Z_1$ mit $(w_2)\in Z_2$ ist $[\infty]$ selbst.

(6) Verbindung von (z) und (∞).

Die Verbindungsgerade ist $[\infty]$ und eine Umgebung von ihr sei wie in (5) gegeben. Man muß eine Umgebung Z von (z) und eine Umgebung W von (∞) finden, so daß die Verbindungsgerade eines Punktes (x, y) aus Z oder eines Punktes (s) aus Z mit einem Punkt (u, v) aus W oder (w) aus W stets in der betrachteten Umgebung von $[\infty]$ liegt. Ist $x=u$, so ist $[u]$ die Verbindungsgerade von (x, y) und (u, v), und es ist $u^{-1}\in U$, wenn man nur Z genügend klein gewählt hat. Für $x\neq u$ bestimmt sich die Verbindungsgerade $[s, t]$ aus $y=sx+t$ und $v=su+t$. Ist $u=0$ und sind Z sowie W hinreichend klein, so ist $t^{-1}\in U$ und $t^{-1}s=t^{-1}(yx^{-1})-x^{-1}\in U$. Ist $y=v$, so ist $s=0$ und für hinreichend kleines Z und W ist wieder $t^{-1}\in U$, $t^{-1}s=0\in U$. Ist $y\neq v$, so ist

$s \neq 0$. Wäre $t = 0$, so wäre $(y x^{-1})(u v^{-1}) = 1$. Man kann Z und W von vornherein so klein gewählt denken, daß dies nicht der Fall sein kann. Dann muß also auch $t \neq 0$ sein. Jetzt kann man $y = s x + t$ und $v = s u + t$ umformen zu $t^{-1}(y x^{-1}) = t^{-1} s + x^{-1}$ und $t^{-1} = (t^{-1} s)(u v^{-1}) + v^{-1}$. Daraus ergibt sich durch Elimination von $t^{-1} s$ die Gleichung

$$t^{-1} - \left(t^{-1}(y x^{-1})\right)(u v^{-1}) = v^{-1} - x^{-1}(u v^{-1}).$$

Mit den Abkürzungen $a = y x^{-1}$, $b = u v^{-1}$, $c = v^{-1} - x^{-1}(u v^{-1})$ liegen für hinreichend kleine Umgebungen Z und W die Elemente b, c in einer beliebig kleinen Umgebung von 0. Da der Fall $u = 0$ schon behandelt ist, kann man sich auf $u \neq 0$ beschränken, also $b \neq 0$ annehmen. Z und W waren darüber hinaus so gewählt, daß $b^{-1} \neq a$ ist. Nun ergibt sich

$$t^{-1} = (c b^{-1})(b^{-1} - a)^{-1} = c + (c a)(b^{-1} - a)^{-1} = c + (c a)\left(b(1 - a b)^{-1}\right).$$

Man kann also durch Wahl von Z und W erreichen, daß $t^{-1} \in U$, und ebenso, daß $t^{-1} s = t^{-1}(y x^{-1}) - x^{-1} \in U$.

Für (x, y) aus Z und (w) aus W bestimmt sich die Verbindungsgerade $[w, t]$ aus $y = w x + t$. Dabei kann man $t \neq 0$ annehmen; denn $t = 0$ hat $w^{-1}(y x^{-1}) = 1$ zur Folge und man kann sich W und Z von vornherein so klein gewählt denken, daß dies nicht der Fall ist. Dann ist $t^{-1} = -x^{-1}\left(w^{-1}(1 - (y x^{-1}) w^{-1})^{-1}\right)$ und liegt für genügend kleine Umgebungen W und Z in U. Dasselbe gilt auch für $t^{-1} w = t^{-1}(y x^{-1}) - x^{-1}$.

Für (s) aus Z und (u, v) aus W erhält man die Verbindungsgerade $[s, t]$ aus $v = s u + t$. Für hinreichend kleine Umgebungen W und Z ist $t \neq 0$. Im Fall $u \neq 0$ ist $(t^{-1} s)(u v^{-1}) + v^{-1} = t^{-1}$ und $t^{-1} = v^{-1} + (v^{-1} s)(v u^{-1} - s)^{-1} = v^{-1} + (v^{-1} s)\left((u v^{-1})(1 - s(u v^{-1}))^{-1}\right)$. Im Fall $u = 0$ ist $t^{-1} = v^{-1}$. Wählt man W und Z hinreichend klein, so ist also $t^{-1} \in U$ und $t^{-1} s \in U$.

Damit ist in jedem der möglichen Fälle gezeigt, daß die Verbindungsgerade zweier Punkte stetig von den beiden Punkten abhängt. Nach der vorausgeschickten Bemerkung gilt auch die duale Aussage. Die angegebene Konstruktion liefert also wirklich eine topologische projektive Ebene.

Literatur

ALEXANDROFF-HOPF:

[**1935**] Topologie. Berlin.

N. BOURBAKI:

[**1947**] Topologie générale V. Paris.

[**1951**] Topologie générale I, 2. Aufl. Paris.

P. ERDÖS:

[**1940**] The dimension of the rational points in Hilbert space. Ann. of Math. **41**, 734—736; MRev. **2**, 178; Zbl. **25**, 187.

H. FREUDENTHAL:

[**1957**] Kompakte projektive Ebenen. Illinois J. of Math. **1**, 9—13.

HUREWICZ-WALLMANN:

[**1948**] Dimension theory. Princeton.

C. KURATOWSKI:
[1952] Topologie II, 2. Aufl. Warschau.
F. R. MOULTON:
[1902] A simple non-desargian plane geometry. Trans. Amer. Math. Soc. **3**, 192—195; FdM **33**, 497.
L. J. PAIGE:
[1950] Neo fields. Duke Math. J. **16**, 39—60; Zbl. **40**, 305.
G. PICKERT:
[1955] Projektive Ebenen. Berlin-Göttingen-Heidelberg.
[1956] Projektive Ebenen über Neokörpern. Wiss. Z. Univ. Jena **1955/56**, 131—135.
L. PONTRJAGIN:
[1946] Topological groups. Princeton.
H. SALZMANN:
[1955] Über den Zusammenhang in topologischen projektiven Ebenen. Math. Z. **61**, 489—494; Zbl. **64**, 178.
L. A. SKORNJAKOV:
[1954] Topologische projektive Ebenen. [Russ.] Trudy Moskov. Mat. Obšč. **3**, 347—373; MRev. **16**, 60.

Mathematisches Seminar der Universität Frankfurt a. M.

(Eingegangen am 24. Dezember 1956)

WESTON, J. D.
Math. Zeitschr. Bd. 67, S. 467—473 (1957)

On integration with respect to a measure

By

J. D. WESTON

1. There are various methods of constructing a Lebesgue integral with respect to an abstract measure (see, for example, [1], Ch. I, [2], Ch. XI, [3], Ch. V; also, as regards the extension of integrals, [4], Ch. IV, [5], Ch. III, [6], Ch. 9). This variety suggests that there is a need for a simple axiomatic definition of an integral with respect to a given measure, and for a proof of uniqueness. We shall exhibit a set of axioms, representing some very elementary properties of an integral, and we shall show that there is a largest class of measurable functions for which an integral having these properties exists; we shall also show that such an integral is uniquely determined by the given measure, and we shall establish the usual convergence properties. From these results it follows immediately that all the definitions of an abstract Lebesgue integral that have been given, or that could be given, are equivalent.

Our method is based on the idea of a continuous measure ("absolutely continuous additive function of a set"). This, with our particular choice of axioms, enables us to define dominated sequences at an early stage, and the main results are then obtained very rapidly by means of EGOROFF's theorem. For convenience, we restrict the discussion to the case of measures which are totally σ-finite ([3], p. 31); this does not involve a serious loss of generality (see, for example, [5], pp. 36—37).

In the next section we state our assumptions (the usual elementary facts) concerning measure and measurable functions. After that, our account of the integration process is given in full detail.

2. We consider an abstract space X and an additive class $\mathscr{A}$ of subsets of X. All "sets" that occur will be members of $\mathscr{A}$. ($\mathscr{A}$ is characterized by the fact that it contains X itself, the complement $X \setminus E$ of any member E, and the union $\bigcup_k E_k$ of any sequence of members $\{E_k\}$; hence $\mathscr{A}$ also contains the intersection $\bigcap_k E_k$ of any sequence of members.)

By a *measurable function* we mean a real-valued function f on X such that $f^{-1}(B) \in \mathscr{A}$ for every BOREL set B of real numbers. For example, the characteristic function χ_E of any set $E \in \mathscr{A}$ is measurable; more generally, a function is measurable if it is *simple* in the sense that its value is constant on each of a finite number of sets which belong to $\mathscr{A}$ and compose X. The limit of a convergent sequence of measurable functions is measurable, and every measurable function is the limit of a sequence of simple functions

(whence it follows that sums and products of measurable functions are measurable); in fact if f is measurable there is a sequence of simple functions $\{g_n\}$ such that, for each $x \in X$, $|g_n(x)| \leq |f(x)|$ for all n and $g_n(x) \to f(x)$ as $n \to \infty$. If f is measurable then $|f|$ is measurable.

By a *measure* we mean a function μ on $\mathscr{A}$ with $0 \leq \mu(E) \leq \infty$ for every $E \in \mathscr{A}$ and $\mu\left(\bigcup_k E_k\right) = \sum_k \mu(E_k)$ for every sequence $\{E_k\}$ in $\mathscr{A}$ such that $E_i \cap E_j = 0$ when $i \neq j$. If $\{E_k\}$ is a sequence in $\mathscr{A}$ such that $E_k \subseteq E_{k+1}$ for every k, then $\mu\left(\bigcup_k E_k\right) = \lim_{k \to \infty} \mu(E_k)$; and if $E_{k+1} \subseteq E_k$ then $\mu\left(\bigcap_k E_k\right) = \lim_{k \to \infty} \mu(E_k)$ provided that $\mu(E_k) < \infty$ for some k. The abbreviation "*p.p.*(μ)" means "except in a set H for which $\mu(H) = 0$".

We shall use the following version of EGOROFF'S theorem: *Let μ be a measure on $\mathscr{A}$, and let $\{f_n\}$ be a sequence, or a multiple sequence, of measurable functions converging p.p.(μ) in a set E for which $\mu(E) < \infty$. Then if $\varepsilon > 0$ there is a subset H of E such that $\mu(H) < \varepsilon$ and $\{f_n\}$ converges uniformly in $E \backslash H$.*

3. We suppose that μ is a measure such that X can be expressed as $\bigcup_k X_k$, where, for each k, $X_k \subseteq X_{k+1}$ and $\mu(X_k) < \infty$; and we consider a class $\mathscr{L}$ of measurable functions such that

(a) $\chi_E \in \mathscr{L}$ if $\mu(E) < \infty$,

(b) $\alpha f + \beta g \in \mathscr{L}$ if $f, g \in \mathscr{L}$ and α, β are constants,

(c) $|f| \in \mathscr{L}$ if $f \in \mathscr{L}$.

There is a smallest class, $\mathscr{L}_0$, having the properties (a), (b), (c); this consists of all functions of the type $\sum_{i=1}^{n} \alpha_i \chi_{E_i}$, where $\mu(E_i) < \infty$ for each i and $E_i \cap E_j = 0$ when $i \neq j$. Every measurable function f is the limit of a sequence $\{h_n\}$ in $\mathscr{L}_0$; for if $\{g_n\}$ is a sequence of simple functions converging to f, and $h_n = g_n \chi_{X_n}$, then $h_n \in \mathscr{L}_0$ for each n, and $\{h_n\}$ converges to f. Moreover, we may suppose that $|h_n(x)| \leq |f(x)|$ for all x and all n (since $|h_n(x)| \leq |g_n(x)|$).

By an *integral* (with respect to μ) we mean a real-valued function $\int$ on $\mathscr{A} \times \mathscr{L}$, subject to the following conditions (we use an obvious notation):

(1) If $f \in \mathscr{L}$ and f has a constant value c on E, where $\mu(E) < \infty$, then

$$\int_E f = c\mu(E).$$

(2) For each $E \in \mathscr{A}$, $\int_E$ is a linear form on $\mathscr{L}$; that is,

$$\int_E (\alpha f + \beta g) = \alpha \int_E f + \beta \int_E g.$$

(3) For each $f \in \mathscr{L}$, $\int |f|$ is a measure on $\mathscr{A}$.

It follows from (2) and (3) that, for each $f \in \mathscr{L}$, $\int f$ is an additive function on $\mathscr{A}$. Indeed, if $g = \frac{1}{2}(|f| + f)$ and $h = \frac{1}{2}(|f| - f)$ then $g, h \in \mathscr{L}$ and $\int g$, $\int h$ are measures (with finite values); hence if $\{E_k\}$ is a sequence in $\mathscr{A}$ such that

$E_i \cap E_j = 0$ when $i \neq j$ then, since $f = g - h$,

$$\int_{\bigcup_k E_k} f = \int_{\bigcup_k E_k} g - \int_{\bigcup_k E_k} h = \sum_k \int_{E_k} g - \sum_k \int_{E_k} h = \sum_k \int_{E_k} f,$$

all the series being absolutely convergent.

An integral can be defined on $\mathscr{A} \times \mathscr{L}_0$ in the following way: if $E \in \mathscr{A}$ and $f = \sum_{i=1}^{n} \alpha_i \chi_{E_i}$, where $\mu(E_i) < \infty$ for each i and $E_i \cap E_j = 0$ when $i \neq j$, let

$$\int_E f = \sum_{i=1}^{n} \alpha_i \mu(E \cap E_i).$$

There is no ambiguity here, for the sum on the right is not altered if we take $\alpha_1, \ldots, \alpha_n$ to be the distinct values of f. It is easy to verify that the conditions (1), (2), (3) are satisfied.

In all that follows we assume $\int$ to be some fixed integral on $\mathscr{A} \times \mathscr{L}$. (We eventually obtain a necessary and sufficient condition on $\mathscr{L}$ in order that this integral may exist, and we prove that it is unique.)

THEOREM 1. *Suppose that $f(x) \leq g(x)$ for all $x \in E$, where $E \in \mathscr{A}$ and $f, g \in \mathscr{L}$. Then $\int_E f \leq \int_E g$.*

Proof. Let $h = |g - f| - (g - f)$. Then $h \in \mathscr{L}$ and $h = |h|$. Hence, since $\int |h|$ is a measure,

$$\int_E h = \lim_{k \to \infty} \int_{E \cap X_k} h.$$

But $h(x) = 0$ for all $x \in E$; so that, since $\mu(E \cap X_k) < \infty$, $\int_{E \cap X_k} h = 0$ for each k, by (1). Thus $\int_E h = 0$. By (2), therefore,

$$\int_E |g - f| - \int_E g + \int_E f = \int_E h = 0.$$

Since $\int_E |g - f| \geq 0$, it follows that $\int_E f \leq \int_E g$.

Theorem 1 has the following corollaries (for $E \in \mathscr{A}$ and $f, g \in \mathscr{L}$):

(i) When $\mu(E) < \infty$, $\int_E f \leq c\mu(E)$ if $f(x) \leq c$ for all $x \in E$ [since, by (1), $\int_E g = c\mu(E)$ if $g = c\chi_E$]; and $\int_E f \geq c\mu(E)$ if $f(x) \geq c$ for all $x \in E$.

(ii) $\int_E f = \int_E g$ if $f(x) = g(x)$ for all $x \in E$; in particular, $\int_E f = 0$ if $f(x) = 0$ for all $x \in E$ [since, by (2), $\int_E g = 0$ if $g(x) = 0$ for all x].

(iii) $\left| \int_E f \right| \leq \int_E |f|$ [since $\pm f(x) \leq |f(x)|$ for all x].

THEOREM 2. *In order that $\int_E |f| = 0$, where $E \in \mathscr{A}$ and $f \in \mathscr{L}$, it is necessary and sufficient that $f(x) = 0$ p.p.(μ) in E.*

Proof. Suppose first that $\int_E |f| = 0$. For each k, let E_k be the set of all points $x \in E \cap X_k$ for which $|f(x)| \geq 1/k$. Then $E_k \in \mathscr{A}$ and $\mu(E_k) < \infty$, so that

$$(1/k)\,\mu(E_k) \leq \int_{E_k} |f| \leq \int_E |f| = 0,$$

and therefore $\mu(E_k) = 0$, for each k. Hence $\mu\left(\bigcup_k E_k\right) = 0$. But $\bigcup_k E_k$ is the set of all points $x \in E$ for which $f(x) \neq 0$; thus $f(x) = 0$ $p.p.(\mu)$ in E.

On the other hand, suppose that $f(x) = 0$ $p.p.(\mu)$ in E. For each k, let H_k be the set of all points $x \in E$ for which $0 < |f(x)| < k$. Then $H_k \in \mathscr{A}$, $H_k \subseteq H_{k+1}$, and $\mu(H_k) = 0$. Hence if $H = \bigcup_k H_k$ then

$$\int_H |f| = \lim_{k \to \infty} \int_{H_k} |f| \leq \lim_{k \to \infty} k\,\mu(H_k) = 0.$$

But $f(x) = 0$ for all $x \in E \backslash H$, so that $\int_{E \backslash H} |f| = 0$. Since $E = (E \backslash H) \cup H$, it follows that $\int_E |f| = 0$.

Theorem 2 enables us to relax the hypothesis of Theorem 1: it is enough to have $f(x) \leq g(x)$ $p.p.(\mu)$ in E.

4. A measure ν is *continuous*[1]) (with respect to μ) if $\nu(X) < \infty$ and $\nu(E) = 0$ when $\mu(E) = 0$. Theorem 2 shows that, for each $f \in \mathscr{L}$, the measure $\int |f|$ is continuous [for if $\mu(E) = 0$ then $f(x) = 0$ $p.p.(\mu)$ in E]. Henceforth we shall write "$p.p.$" to mean "$p.p.(\mu)$", and hence also "$p.p.(\nu)$" for any continuous measure ν.

We say that a sequence $\{f_n\}$ in $\mathscr{L}$ is *dominated* (with respect to $\int$) if there is a continuous measure ν such that $\int_E |f_n| \leq \nu(E)$ for all $E \in \mathscr{A}$ and all n. For example, we might have $|f_n(x)| \leq |f(x)|$ for all x and all n, where $f \in \mathscr{L}$: the measure $\int |f|$ then dominates $\{f_n\}$.

LEMMA. *Let $\{f_n\}$ and $\{g_n\}$ be dominated sequences in $\mathscr{L}$ such that $f_m(x) - g_n(x) \to 0$ $p.p.$ in E as $m, n \to \infty$, where $E \in \mathscr{A}$. Then $\int_E f_m - \int_E g_n \to 0$ as $m, n \to \infty$.*

Proof. Suppose that ν' and ν'' dominate $\{f_n\}$ and $\{g_n\}$ respectively, and let $\nu = \nu' + \nu''$. Then ν is a continuous measure that dominates both sequences. If $\varepsilon > 0$, we can choose k so that $\nu(X \backslash X_k) < \varepsilon$ [since $\nu(X) = \lim_{k \to \infty} \nu(X_k)$ and $\nu(X) < \infty$]. Then, for all m and n,

$$\int_{E \backslash X_k} |f_m - g_n| \leq \int_{E \backslash X_k} (|f_m| + |g_n|) \leq 2\nu(E \backslash X_k) \leq 2\nu(X \backslash X_k) < 2\varepsilon.$$

By EGOROFF's theorem (with ν in place of μ), $E \cap X_k$ has a subset H such that $\nu(H) < \varepsilon$ and the double sequence $\{f_m - g_n\}$ is uniformly convergent in $(E \cap X_k) \backslash H = E_k$, say. Thus

$$\int_H |f_m - g_n| \leq 2\nu(H) \leq 2\varepsilon$$

[1]) A discussion of continuous measures may be found in [1], pp. 30—36; here we require only the properties specified in the definition.

for all m and n, and if m and n are large enough then

$$|f_m(x) - g_n(x)| \leq \varepsilon/\mu(E_k), \qquad \text{all } x \in E_k,$$

so that $\int_{E_k} |f_m - g_n| \leq \varepsilon$. Since $E = E_k \cup H \cup (E \setminus X_k)$, it follows that

$$\left| \int_E f_m - \int_E g_n \right| \leq \int_E |f_m - g_n| < 5\varepsilon$$

when m and n are large enough; thus $\int_E f_m - \int_E g_n \to 0$ as $m, n \to \infty$.

THEOREM 3. *Let $\{f_n\}$ be a dominated sequence in $\mathscr{L}$ such that $f_n(x) \to f(x)$ p.p. in E as $n \to \infty$, where $f \in \mathscr{L}$ and $E \in \mathscr{A}$. Then $\int_E f_n \to \int_E f$ as $n \to \infty$.*

Proof. Let $g_n = f$ for every n. Then $\int |f|$ dominates $\{g_n\}$. Hence, by the lemma, $\int_E f_n - \int_E f \to 0$ as $n \to \infty$, so that $\int_E f_n \to \int_E f$.

5. Suppose that $f = \sum_{i=1}^{n} \alpha_i \chi_{E_i}$, where $\mu(E_i) < \infty$ for each i (so that $f \in \mathscr{L}_0$). If $E \in \mathscr{A}$ then, for each i, $\int_{E \cap E_i} \chi_{E_i} = \mu(E \cap E_i)$ by (1), and $\int_{E \setminus E_i} \chi_{E_i} = 0$ since $\chi_{E_i}(x) = 0$ for all $x \in E \setminus E_i$; thus $\int_E \chi_{E_i} = \mu(E \cap E_i)$. Hence, by (2),

$$\int_E f = \sum_{i=1}^{n} \alpha_i \mu(E \cap E_i).$$

This shows that there is only one integral on $\mathscr{A} \times \mathscr{L}_0$. As far as $\mathscr{L}_0$ is concerned, therefore, the notion of dominated sequence is uniquely defined.

We now define $\mathscr{L}_1$ to be the class of all measurable functions which are limits p.p. of dominated sequences in $\mathscr{L}_0$. It is easy to verify that $\mathscr{L}_1$ has the properties (a), (b), (c). [For (a), use the fact that a sequence in which every term is χ_E is dominated by $\int \chi_E$, and for (b) the fact that if ν' and ν'' dominate $\{g_n\}$ and $\{h_n\}$ respectively then $|\alpha|\nu' + |\beta|\nu''$ dominates $\{\alpha g_n + \beta h_n\}$; (c) is obvious.]

THEOREM 4. *There is an integral on $\mathscr{A} \times \mathscr{L}_1$.*

Proof. Let $f \in \mathscr{L}_1$, and let $\{h_n\}$ be a dominated sequence in $\mathscr{L}_0$ such that $h_n(x) \to f(x)$ p.p. as $n \to \infty$. Then $h_m(x) - h_n(x) \to 0$ p.p. as $m, n \to \infty$. Hence if $E \in \mathscr{A}$ then, by the lemma, $\int_E h_m - \int_E h_n \to 0$ as $m, n \to \infty$, so that $\left\{ \int_E h_n \right\}$ is a convergent sequence. Let $\overline{\int}_E f = \lim_{n \to \infty} \int_E h_n$; this is independent of the choice of $\{h_n\}$, for if $\{h_n'\}$ is an alternative choice then, by the lemma, $\int_E h_n - \int_E h_n' \to 0$ as $n \to \infty$.

With this definition of $\overline{\int}$, (2) is obviously satisfied and, since we may suppose that $|h_n(x)| \leq |f(x)|$ for all x and all n, (1) is satisfied when $c = 0$; consideration of $f - c\chi_E$ then shows that (1) is satisfied for any value of c. As to (3), let $\{E_k\}$ be a sequence in $\mathscr{A}$ such that $E_i \cap E_j = 0$ when $i \neq j$. The series $\sum_k \int_{E_k} |h_n|$

converges uniformly with respect to n, by comparison with $\sum_k \nu(E_k)$, where ν dominates $\{h_n\}$. Hence

$$\overline{\int}_{\bigcup_k E_k} |f| = \lim_{n\to\infty} \int_{\bigcup_k E_k} |h_n| = \lim_{n\to\infty} \sum_k \int_{E_k} |h_n| = \sum_k \overline{\int}_{E_k} |f|,$$

so that $\overline{\int}|f|$ is a measure (being obviously non-negative). Thus we have defined an integral $\overline{\int}$ on $\mathscr{A}\times\mathscr{L}_1$.

If $\{f_n\}$ is a sequence of measurable functions such that $\{f_n(x)\}$ converges $p.p.$, let f be defined as follows: $f(x)=\lim_{n\to\infty} f_n(x)$ if $\{f_n(x)\}$ converges, and $f(x)=0$ otherwise. Then f is a measurable function, which we denote by "$\lim_{n\to\infty} f_n$".

THEOREM 5. *Let $\{f_n\}$ be a dominated sequence in $\mathscr{L}$ such that $\{f_n(x)\}$ converges $p.p.$ Then $\lim_{n\to\infty} f_n \in \mathscr{L}_1$.*

Proof. For each n, there is a sequence $\{h_m^{(n)}\}$ in $\mathscr{L}_0$ such that, for each x, $|h_m^{(n)}(x)| \le |f_n(x)|$ for all m and $h_m^{(n)}(x)\to f_n(x)$ as $m\to\infty$. Suppose that ν dominates $\{f_n\}$. Since $\nu(X)<\infty$, EGOROFF's theorem shows that, for each n, there is a set $H_n\in\mathscr{A}$ such that $\nu(H_n)<1/2^n$ and $h_m^{(n)}(x)\to f_n(x)$ uniformly in $X\backslash H_n$, as $m\to\infty$; thus we can choose m_n so that, for all $x\in X\backslash H_n$,

$$|h_{m_n}^{(n)}(x) - f_n(x)| < 1/n.$$

Let $K_n=\bigcup_{p>n} H_p$, and let $K=\bigcap_n K_n$. Then $K\in\mathscr{A}$ and, since $\nu(K_n)<1/2^n$, $\nu(K)=0$. Let $h_n=h_{m_n}^{(n)}$, for each n. If $x\in X\backslash K$ then $x\in X\backslash K_n$ for some n, so that $x\in X\backslash H_n$ when n is large enough. Thus $h_n(x)-f_n(x)\to 0$ as $n\to\infty$, for each $x\in X\backslash K$. But if $f=\lim_{n\to\infty} f_n$ then $f_n(x)\to f(x)$ $p.p.$; hence $h_n(x)\to f(x)$ $p.p.$ in $X\backslash K$, as $n\to\infty$.

Now for each n, $\int_K |f_n| = 0$ [since $\nu(K)=0$] and therefore, by Theorem 2, $f_n(x)=0$ for all $x\in K\backslash L_n$, where L_n is a subset of K for which $\mu(L_n)=0$. Let $L=\bigcup_n L_n$. Then $\mu(L)=0$, and if $x\in K\backslash L$ then $f_n(x)=0$, and therefore $h_n(x)=0$, for all n. Thus $f(x)=0$, and $h_n(x)\to 0$, $p.p.$ in K. Hence $h_n(x)\to f(x)$ $p.p.$ as $n\to\infty$. Since $\{h_n\}$ is dominated by ν, it follows that $f\in\mathscr{L}_1$.

THEOREM 6. *The integral $\int$ on $\mathscr{A}\times\mathscr{L}$ is unique, and $\mathscr{L}\subseteq\mathscr{L}_1$.*

Proof. Let $f\in\mathscr{L}_1$, and let $f_n=f$ for every n. Then $\int|f|$ dominates $\{f_n\}$, so that, by Theorem 5, $f\in\mathscr{L}_1$; thus $\mathscr{L}\subseteq\mathscr{L}_1$. By the definition of $\mathscr{L}_1$, there is a dominated sequence $\{h_n\}$ in $\mathscr{L}_0$ such that $h_n(x)\to f(x)$ $p.p.$ as $n\to\infty$. By Theorem 3, $\int_E h_n \to \int_E f$ as $n\to\infty$, for any $E\in\mathscr{A}$. Since the integral on $\mathscr{A}\times\mathscr{L}_0$ is unique, it follows that $\int$ is unique on $\mathscr{A}\times\mathscr{L}$ (any integral on $\mathscr{A}\times\mathscr{L}$ coincides with $\int$ on $\mathscr{A}\times\mathscr{L}_0$ and hence also on $\mathscr{A}\times\mathscr{L}$).

We have now shown that for an integral to exist on $\mathscr{A}\times\mathscr{L}$ it is necessary and sufficient that $\mathscr{L}\subseteq\mathscr{L}_1$. It is possible to have $\mathscr{L}_0\subset\mathscr{L}\subset\mathscr{L}_1$: for example, if X is the interval $(0, 1)$ and μ is Lebesgue measure, $\mathscr{L}$ might consist of all bounded measurable functions.

6. To conclude, we take $\mathscr{L} = \mathscr{L}_1$ and prove two theorems that help to elucidate the structure of this class.

THEOREM 7. *If f is a measurable function such that $|f(x)| \leq g(x)$ p.p., where $g \in \mathscr{L}_1$, then $f \in \mathscr{L}_1$. In particular, $f \in \mathscr{L}_1$ if $|f| \in \mathscr{L}_1$ and f is measurable.*

Proof. Let $\{h_n\}$ be a sequence in $\mathscr{L}_0$ such that, for each x, $|h_n(x)| \leq |f(x)|$ for all n and $h_n(x) \to f(x)$ as $n \to \infty$. If $|f(x)| \leq g(x)$ p.p. and $g \in \mathscr{L}_1$ then $\int g$ dominates $\{h_n\}$; hence $f \in \mathscr{L}_1$.

The results we have obtained so far evidently include LEBESGUE'S theorem of dominated convergence: *if $\{f_n\}$ is a sequence of measurable functions such that $\{f_n(x)\}$ converges p.p. and, for all n, $|f_n(x)| \leq g(x)$ p.p., where $g \in \mathscr{L}_1$, then for each $E \in \mathscr{A}$,*

$$\int_E \lim_{n\to\infty} f_n = \lim_{n\to\infty} \int_E f_n .$$

LEBESGUE'S other convergence theorem may be deduced from the following result.

THEOREM 8. *Let $\{f_n\}$ be a non-decreasing sequence of non-negative functions in $\mathscr{L}_1$ such that the sequence $\left\{\int_X f_n\right\}$ is bounded. Then $\{f_n(x)\}$ converges p.p., and $\{f_n\}$ is dominated.*

Proof. Let $c = \sup_n \int_X f_n$. For each m and n, let $H_{m,n}$ be the set of all points x for which $f_n(x) > m$. Then $H_{m,n} \in \mathscr{A}$, $H_{m,n} \subseteq H_{m,n+1}$, and $\mu(H_{m,n}) \leq c/m$. Let $H_m = \bigcup_n H_{m,n}$. Then $\mu(H_m) = \lim_{n\to\infty} \mu(H_{m,n}) \leq c/m$. But H_m is the set of all points x for which $\lim_{n\to\infty} f_n(x) > m$, so that $H_{m+1} \subseteq H_m$ and $\bigcap_m H_m$ is the set in which $\{f_n(x)\}$ diverges. Hence $\mu\left(\bigcap_m H_m\right) = 0$, so that $\{f_n(x)\}$ converges p.p.

Let $\nu_n = \int |f_n|$ for each n. Then, for each $E \in \mathscr{A}$, $\{\nu_n(E)\}$ is a bounded non-decreasing sequence: let $\nu(E) = \lim_{n\to\infty} \nu_n(E)$. If $\{E_k\}$ is a sequence in $\mathscr{A}$ such that $E_i \cap E_j = 0$ when $i \neq j$ then, for each k,

$$\sum_{i=1}^{k} \nu(E_i) = \lim_{n\to\infty} \sum_{i=1}^{k} \nu_n(E) \leq \lim_{n\to\infty} \nu_n\Big(\bigcup_k E_k\Big) = \nu\Big(\bigcup_k E_k\Big).$$

Hence $\sum_k \nu(E_k) \leq \nu\left(\bigcup_k E_k\right)$. But if $\varepsilon > 0$ we can choose n so that

$$\nu\Big(\bigcup_k E_k\Big) \leq \nu_n\Big(\bigcup_k E_k\Big) + \varepsilon = \sum_k \nu_n(E_k) + \varepsilon \leq \sum_k \nu(E_k) + \varepsilon .$$

Thus $\nu\left(\bigcup_k E_k\right) = \sum_k \nu(E_k)$, so that ν is a measure. Obviously, ν is continuous and dominates $\{f_n\}$.

References

[1] SAKS, S.: Theory of the integral. Warsaw 1937. — [2] NEUMANN, J. v.: Functional operators I. Princeton 1950. — [3] HALMOS, P. R.: Measure theory. New York 1950. — [4] BOURBAKI, N.: Éléments de mathématique XIII. Paris 1952. — [5] LOOMIS, L. H.: An introduction to abstract harmonic analysis. New York 1953. — [6] AUMANN, G.: Reelle Funktionen. Berlin 1954.

University of Durham, Newcastle upon Tyne

(Eingegangen am 7. Januar 1957)

MOPPERT, K.-F.
Math. Zeitschr. Bd. 67, S. 474—478 (1957)

Funktionenscharen im L_2

Von

KARL-FELIX MOPPERT *

Wir beschäftigen uns im folgenden mit reellen Funktionen $f(\nu, t)$ der reellen Variablen ν und t, die als Funktionen von t dem reellen Hilbertschen Raum L_2, bezogen auf das Intervall $0 \leq t \leq 1$ angehören. Norm bzw. Skalarprodukt in diesem Raum bezeichnen wir mit $\| \|$ bzw. $(,)$.

Definition 1. Ist $f(\nu, t) \in L_2$ für $N_1 < \nu < N_2$ und ist für $N_1 < \nu_\varkappa < N_2$, $\lim_{\varkappa \to \infty} \nu_\varkappa = N_1$ $(\varkappa = 1, 2, \ldots)$ die Funktionenschar $f(\nu_\varkappa, t)$ abgeschlossen, so nennen wir $f(\nu, t)$ bei $\nu = N_1$ *rechtsseitig unbedingt abgeschlossen*. Analog definieren wir *„linksseitig unbedingt abgeschlossen"* bzw. *„unbedingt abgeschlossen"* schlechthin.

Bekanntlich gibt es analytische Funktionen $f(\nu, t)$, die bei einem $\nu = N_1$ nicht unbedingt abgeschlossen sind, während es doch gewisse gegen N_1 strebende Zahlenfolgen $\nu_\varkappa^*$ $(\varkappa = 1, 2, \ldots)$ gibt, derart, daß die Funktionen $f(\nu_\varkappa^*, t)$ ein abgeschlossenes System bilden. So ist z.B. $f(\nu, t) = t^\nu$ bei jedem N_1, $-\frac{1}{2} < N_1 < \infty$ unbedingt abgeschlossen, bei $\nu = -\frac{1}{2}$ aber nicht rechtsseitig unbedingt abgeschlossen, während doch die Funktionenschar $f_\varkappa(t) = t^{-\frac{1}{2} + \frac{1}{\varkappa}}$ $(\varkappa = 1, 2, \ldots)$ abgeschlossen ist [1]).

Wir geben in dieser Arbeit zunächst eine hinreichende Bedingung dafür, daß $f(\nu, t)$ bei $\nu = 0$ rechtsseitig unbedingt abgeschlossen ist und hierauf, unter etwas stärkeren Voraussetzungen, auch eine notwendige. Die Ergebnisse lassen sich natürlich ohne weiteres so formulieren, daß sie sich auf eine der beiden andern Formen der unbedingten Abgeschlossenheit bei einem beliebigen $\nu = N_1$ beziehen.

§ 1. Hinreichende Bedingung dafür, daß $f(\nu, t)$ bei $\nu = 0$ rechtsseitig unbedingt abgeschlossen ist

Wir bezeichnen mit $G(f_1, f_2, \ldots, f_n)$ die Gramsche Determinante der Funktionen $f_\nu(t) \in L_2$; $\nu = 1, 2, \ldots, n$, bezogen auf das Intervall $(0, 1)$. Wir setzen weiter $Q(g; f_1, f_2, \ldots, f_n) = G(g, f_1, f_2, \ldots, f_n)/G(f_1, f_2, \ldots, f_n)$. Abkürzend schreiben wir f_{ν_i} für $f(\nu_i, t)$ und $f^{(\varkappa)}$ für $\left[\left(\frac{\partial}{\partial \nu}\right)^\varkappa f(\nu, t)\right]_{\nu = 0}$. Wir definieren das Symbol $[\nu_1 \nu_2 \ldots \nu_k] f$ $(k = 1, 2, \ldots)$ durch $[\nu_1] f = f_{\nu_1}$ und $[\nu_1 \nu_2 \ldots \nu_{k+1}] f = \left([\nu_1 \nu_2 \ldots \nu_{k-1} \nu_{k+1}] f - [\nu_1 \nu_2 \ldots \nu_{k-1} \nu_k] f\right) : (\nu_{k+1} - \nu_k)$ $(k = 1, 2, \ldots)$ [2]).

*) Herrn Dr. H. F. J. LOWIG bin ich für mehrere wertvolle Hinweise zu Dank verpflichtet.

[1]) Vgl. KACZMARZ, S., u. H. STEINHAUS: Theorie der Orthogonalreihen. Monografje Matematyczne, 2. Aufl., Bd. VI, S. 90. New York: Chelsea 1951. Siehe auch die dortigen Literaturangaben.

[2]) Vgl. NÖRLUND: Differenzenrechnung, Grundlehren-Bd. XIII, S. 8. Berlin: Springer 1924.

Hilfssatz. Gilt $\|f_\nu(t)\| \leq M$ und $\|f_\nu(t)+h_\nu(t)\| \leq M$ für $\nu = 1, 2, \ldots, n$, so ist

$$|G(f_1+h_1, f_2+h_2, \ldots, f_n+h_n) - G(f_1, f_2, \ldots, f_n)| \cdot \leq 2M^{2n-1}(\|h_1\| + \|h_2\| + \cdots + \|h_n\|).$$

Beweis. Es kann für jedes $f \in L_2$ eindeutig $f = f^* + f^{**}$ gesetzt werden mit $(f^*, f^{**}) = 0$, f^{**} enthalten in der durch die Funktionen $f_2, f_3, \ldots, f_n$ aufgespannten linearen Mannigfaltigkeit. Dann ist $G(f_1, f_2, \ldots, f_n) = \|f_1^*\|^2 G(f_2, \ldots, f_n)$ und $(f+g)^* = f^* + g^*$. Es folgt

$$\begin{aligned} |G(f_1+h_1, f_2, \ldots, f_n) - G(f_1, f_2, \ldots, f_n)| &= |\|f_1^*+h_1^*\|^2 - \|f_1^*\|^2|\, G(f_2, f_3, \ldots, f_n) \\ &= |\|f_1^*+h_1^*\| - \|f_1^*\||\,(\|f_1^*+h_1^*\| + \|f_1^*\|)\, G(f_2, f_3, \ldots, f_n) \\ &\leq \|h_1^*\|\,(\|f_1^*+h_1^*\| + \|f_1^*\|)\, G(f_2, f_3, \ldots, f_n) \\ &\leq \|h_1\| \cdot 2M^{2n-1} \end{aligned}$$

und hieraus unmittelbar die Behauptung.

Definition 2. Wir nennen $f(\nu, t)$ bei $\nu = \nu_0$ *rechtsseitig k-mal gleichmäßig L_2-differenzierbar*, wenn die folgenden Voraussetzungen erfüllt sind:

a) Es gibt $\varepsilon > 0$ so, daß $\left(\frac{\partial}{\partial \nu}\right)^\varkappa f(h, t)$ existiert und $\in L_2$ ist für $\nu_0 \leq h \leq \nu_0 + \varepsilon$; $\varkappa = 0, 1, \ldots, k$ und für fast alle t in $(0, 1)$.

b) Ist $\nu_0 \leq h(t) \leq \nu_0 + \varepsilon$ und ist $\left(\frac{\partial}{\partial \nu}\right)^\varkappa f(h(t), t) \in L_2$, so ist

$$\left\| \left(\frac{\partial}{\partial \nu}\right)^\varkappa f(h(t), t) - \left(\frac{\partial}{\partial \nu}\right)^\varkappa f(\nu_0, t) \right\| = o(1)$$

für $\varepsilon \to 0$, $\varkappa = 0, 1, \ldots, k$.

Satz 1. Sei $f(\nu, t)$ bei $\nu = 0$ rechtsseitig k-mal gleichmäßig L_2-differenzierbar und sei $G(f^{(0)}, f^{(1)}, \ldots, f^{(k)}) \neq 0$. Dann ist für jedes $g(t) \in L_2$:

$$Q(g; f_{\nu_1}, f_{\nu_2}, \ldots, f_{\nu_{k+1}}) = Q(g; f^{(0)}, f^{(1)}, \ldots, f^{(k)}) + o(1)$$

für $\sqrt{\nu_1^2 + \nu_2^2 + \cdots + \nu_{k+1}^2} \to 0$, sofern bei diesem Grenzübergang die $\nu_\varkappa$ ($\varkappa = 1, 2, \ldots, k+1$) paarweise verschiedene positive Werte durchlaufen.

Beweis. Aus elementaren Determinantensätzen folgt (sofern die Werte $\nu_1, \nu_2, \ldots, \nu_{k+1}$ paarweise verschieden sind), daß wir $\nu_1 < \nu_2 < \cdots < \nu_{k+1}$ voraussetzen können und

$$\begin{aligned} &G(g, f_{\nu_1}, f_{\nu_2}, \ldots, f_{\nu_{k+1}}) \\ &= \prod_{\varkappa=2}^{k+1} (\nu_1 - \nu_\varkappa)^2\, G(g, [\nu_1]f, [\nu_1\nu_2]f, [\nu_1\nu_3]f, \ldots, [\nu_1\nu_{k+1}]f) \\ &= \prod_{\varkappa=2}^{k+1} (\nu_1 - \nu_\varkappa)^2 \prod_{\varkappa=3}^{k+1} (\nu_2 - \nu_\varkappa)^2\, G(g, [\nu_1]f, [\nu_1\nu_2]f, [\nu_1\nu_2\nu_3]f, [\nu_1\nu_2\nu_4]f, \ldots) \\ &= \cdots \\ &= \prod_{\varkappa>\lambda}^{k+1} (\nu_\lambda - \nu_\varkappa)^2\, G(g, [\nu_1]f, [\nu_1\nu_2]f, [\nu_1\nu_2\nu_3]f, \ldots, [\nu_1\nu_2\ldots\nu_{k+1}]f), \end{aligned}$$

wobei im letzten Produkt rechts alle Faktoren der Form $(\nu_\lambda - \nu_\varkappa)^2$ enthalten sind mit $\varkappa > \lambda$, $k+1 \geq \varkappa$, $\lambda \geq 1$.

Die entsprechende Umformung führen wir bei $G(f_{\nu_1}, f_{\nu_2}, \ldots, f_{\nu_{k+1}})$ durch. So folgt

$$Q(g; f_{\nu_1}, f_{\nu_2}, \ldots, f_{\nu_{k+1}}) = Q(g; [\nu_1]f, [\nu_1\nu_2]f, \ldots, [\nu_1\nu_2\ldots\nu_{k+1}]f).$$

Nun gibt es aber[3]) $\nu_\varkappa^*(t)$, $\nu_1 < \nu_\varkappa^*(t) < \nu_{\varkappa+1}$ derart, daß

$$[\nu_1\nu_2\ldots\nu_{\varkappa+1}]f = \frac{1}{\varkappa!}\left(\frac{\partial}{\partial\nu}\right)^\varkappa f(\nu_\varkappa^*(t), t)$$

gilt für $\varkappa = 1, \ldots, k$ und fast alle t in $(0, 1)$. Hier ist der Ausdruck links $\in L_2$ und damit auch derjenige rechts. Somit ist

$$[\nu_1\nu_2\ldots\nu_{\varkappa+1}]f = \frac{1}{\varkappa!}\left(f^{(\varkappa)} + r_\varkappa(t)\right), \quad r_\varkappa(t) \in L_2, \quad \|r_\varkappa(t)\| = o(1)$$

für $\sqrt{\nu_1^2+\nu_2^2+\cdots+\nu_{\varkappa+1}^2} \to 0$ $(\varkappa = 0, 1, \ldots, k)$. Somit ist

$$Q(g; f_{\nu_1}, f_{\nu_2}, \ldots, f_{\nu_{k+1}}) = \frac{G(g, f^{(0)} + r_0(t), f^{(1)} + r_1(t), \ldots, f^{(k)} + r_k(t))}{G(f^{(0)} + r_0(t), f^{(1)} + r_1(t), \ldots, f^{(k)} + r_k(t))}$$

mit $\|r_\varkappa(t)\| = o(1)$ $(\varkappa = 0, 1, \ldots, k)$ für $\sqrt{\nu_1^2+\nu_2^2+\cdots+\nu_{k+1}^2} \to 0$. Anderseits folgt aus dem Hilfssatz:

$$G(g, f^{(0)} + r_0(t), \ldots, f^{(k)} + r_k(t)) = G(g, f^{(0)}, \ldots, f^{(k)}) + R$$

mit $R \to 0$ für $\sqrt{\nu_1^2+\nu_2^2+\cdots+\nu_{k+1}^2} \to 0$. Ebenso ist

$$G(f^{(0)} + r_0(t), \ldots, f^{(k)} + r_k(t)) = G(f^{(0)}, \ldots, f^{(k)}) + R^*$$

mit $R^* \to 0$ für $\sqrt{\nu_1^2+\nu_2^2+\cdots+\nu_{k+1}^2} \to 0$ und damit ist der Satz bewiesen.

Satz 2. Bilden die Funktionen $f^{(0)}, f^{(1)}, \ldots$ ein abgeschlossenes System, ist $G(f^{(0)}, f^{(1)}, \ldots, f^{(\varkappa)}) \not\equiv 0$ für $\varkappa = 1, 2, \ldots$ und ist die Funktion $f(\nu, t)$ bei $\nu = 0$ rechtsseitig beliebig oft gleichmäßig L_2-differenzierbar, so ist $f(\nu, t)$ bei $\nu = 0$ rechtsseitig unbedingt abgeschlossen.

Beweis. Nach Voraussetzung gibt es zu jedem $g(t) \in L_2$ und jedem $\varepsilon > 0$ ein $k_0 = k_0(g, \varepsilon)$ derart, daß $0 \leq Q(g; f^{(0)}, f^{(1)}, \ldots, f^{(k_0-1)}) \leq \frac{\varepsilon}{2}$ ist. Nach Satz 1 gibt es außerdem ein $\delta = \delta(k_0) > 0$ so, daß

$$|Q(g; f_{\nu_1^*}, f_{\nu_2^*}, \ldots, f_{\nu_{k_0}^*}) - Q(g; f^{(0)}, f^{(1)}, \ldots, f^{(k_0-1)})| \leq \frac{\varepsilon}{2}$$

ist für $0 < \nu_\varkappa^* \leq \delta$ $(\varkappa = 1, \ldots, k_0)$, $\nu_\varkappa^* \neq \nu_\lambda^*$ für $\varkappa \neq \lambda$. Ist jetzt $\nu_\varkappa$ $(\varkappa = 0, 1, \ldots)$ eine über positive Werte nach 0 strebende Zahlenfolge, so gibt es $k_1 = k_1(\delta)$ so, daß $0 < \nu_\varkappa \leq \delta$ ist für $\varkappa > k_1$. Dann ist für jedes $k > k_1 + k_0 - 1$:

$$0 \leq Q(g; f_{\nu_1}, f_{\nu_2}, \ldots, f_{\nu_k}) \leq Q(g; f_{\nu_{k_1}}, f_{\nu_{k_1+1}}, \ldots, f_{\nu_{k_1+k_0-1}}) \leq \varepsilon$$

und damit ist der Satz bewiesen.

---

[3]) Vgl. a. a. O. [2]) S. 13.

§ 2. L_2-Potenzreihen

Einen Ausdruck der Form $\sum_0^\infty \nu^\lambda \varphi_\lambda(t)$; $\varphi_\lambda(t) \in L_2$ $(\lambda = 0, 1, \ldots)$ nennen wir eine *L_2-Potenzreihe um $\nu = 0$*. Eine solche hat offenbar ebenso wie eine reelle Potenzreihe einen wohldefinierten Konvergenzradius. Weiter ist eine L_2-Potenzreihe mit Konvergenzradius $r > 0$ für $|\nu| \leq \nu_1 < r$ gleichmäßig konvergent in dem Sinne, daß $\left\| \sum_{\lambda_1}^{\lambda_2} \nu^\lambda \varphi_\lambda(t) \right\| < \varepsilon$ ist für $\lambda_{1,2} \geq l(\varepsilon)$, unabhängig von ν. Somit ist dann auch $f(\nu, t) = \sum_0^\infty \nu^\lambda \varphi_\lambda(t)$ für $|\nu| \leq \nu_1 < r$ beliebig oft nach ν differenzierbar und es ist dort

$$\left(\frac{\partial}{\partial \nu}\right)^\varkappa f(\nu, t) = \sum_\varkappa^\infty \lambda(\lambda - 1) \ldots (\lambda - \varkappa + 1)\, \nu^{\lambda - \varkappa} \varphi_\lambda(t),$$

wobei die rechts stehende Potenzreihe wieder den Konvergenzradius r hat[4]).

Satz 3. Konvergiert $\sum_0^\infty \nu^\lambda \varphi_\lambda(t) = f(\nu, t)$ für $|\nu| < r$, $r > 0$, so ist $f(\nu, t)$ bei $\nu = 0$ beliebig oft gleichmäßig L_2-differenzierbar.

Beweis. Es ist klar, daß Voraussetzung a) von Definition 2 erfüllt ist. Ist jetzt $0 \leq |h(t)| < \varepsilon < \nu_1 < r$, so ist

$$\begin{aligned} \left\| \left(\frac{\partial}{\partial \nu}\right)^\varkappa f(h(t), t) \right\| &= \left\| \sum_\varkappa^\infty \lambda(\lambda - 1) \ldots (\lambda - \varkappa + 1)\, (h(t))^{\lambda - \varkappa} \varphi_\lambda(t) \right\| \\ &\leq \sum_\varkappa^\infty \| \lambda(\lambda - 1) \ldots (\lambda - \varkappa + 1)\, \varepsilon^{\lambda - \varkappa} \varphi_\lambda(t) \| \\ &= \sum_\varkappa^\infty \left(\frac{\varepsilon}{\nu_1}\right)^{\lambda - \varkappa} \| \lambda(\lambda - 1) \ldots (\lambda - \varkappa + 1)\, \nu_1^{\lambda - \varkappa} \varphi_\lambda(t) \| = o(1) \end{aligned}$$

für $\varepsilon \to 0$.

Satz 4. Ist $f(\nu, t)$ für $|\nu| < r$ in eine konvergente L_2-Potenzreihe um $\nu = 0$ entwickelbar: $f(\nu, t) = \sum_0^\infty \nu^\lambda \varphi_\lambda(t)$, $\varphi_\lambda(t) \in L_2$, und ist $G(\varphi_0, \varphi_1, \ldots, \varphi_\varkappa) \not\equiv 0$ für $\varkappa = 1, 2, \ldots$, so sind nur die folgenden Fälle möglich:

a) $f(\nu, t)$ ist bei $\nu = 0$ unbedingt abgeschlossen;

b) es gibt einen echten Teilraum $M \subset L_2$ so, daß $f(\nu, t) \in M$ ist für $|\nu| < r$.

Beweis. Nach dem vorhergehenden Satz ist $f(\nu, t)$ bei $\nu = 0$ beliebig oft gleichmäßig L_2-differenzierbar. Dazu ist $f^{(\varkappa)} = \varkappa!\, \varphi_\varkappa(t)$. Bilden also die Funktionen $\varphi_\lambda(t)$ $(\lambda = 0, 1, \ldots)$ ein abgeschlossenes System, so sind die Voraussetzungen von Satz 2 erfüllt und $f(\nu, t)$ ist bei $\nu = 0$ unbedingt abgeschlossen.

[4]) Vgl. HILLE: Functional Analysis and Semi-Groups. American Mathematical Society Colloquium Publications, Bd. XXXI, S. 52. New York 1948.

Bilden anderseits die Funktionen $\varphi_\lambda(t)$ $(\lambda = 0, 1, \ldots)$ (bzw. die Funktionen $f^{(\varkappa)}$) kein abgeschlossenes System, so gibt es eine Funktion $e(t) \in L_2$ mit $\|e(t)\| = 1$, $(e, f^{(\varkappa)}) = 0$ $(\varkappa = 0, 1, \ldots)$. Sei nun M das orthogonale Komplement von e. Dann kann $f(\nu, t) = \alpha(\nu) e(t) + g(\nu, t)$; $g(\nu, t) \in M$ geschrieben werden mit $\alpha(\nu) = (e, f(\nu, t))$. Anderseits ist aber nach Voraussetzung $f(\nu, t) = \sum_0^\infty \nu^\lambda \varphi_\lambda(t)$. Dann ist für jedes ν mit $|\nu| \leq \nu_0 < r$ wegen der gleichmäßigen Konvergenz der rechtsstehenden Potenzreihe: $\alpha(\nu) = \sum_0^\infty \nu^\lambda (\varphi_\lambda, f(\nu, t))$. Also ist $\alpha(\nu)$ für $|\nu| < r$ in eine konvergente Potenzreihe entwickelbar. Aus $\alpha(\nu) = (e, f(\nu, t))$ folgt aber: $\alpha^{(\varkappa)}(0) = (e, f^{(\varkappa)})$ und dies verschwindet für jedes $\varkappa = 0, 1, \ldots$. Somit ist $f(\nu, t) = g(\nu, t) \in M \subset L_2$ für $|\nu| < r$, w. z. b. w.

University of Tasmania, Hobart (Australien)

(Eingegangen am 1. August 1956)

HUPPERT, B.
Math. Zeitschr. Bd. 67, S. 479—518 (1957)

Lineare auflösbare Gruppen [1]

Von

BERTRAM HUPPERT

Einleitung

Die Untersuchung von endlichen auflösbaren Gruppen von n-reihigen Matrizen (kurz: linearen auflösbaren Gruppen vom Grad n) ist für viele Fragen der abstrakten Gruppentheorie und der Theorie der Permutationsgruppen von Bedeutung.

Ist $\mathfrak{G}$ eine abstrakte auflösbare endliche Gruppe und $\mathfrak{E} = \mathfrak{N}_0 \lhd \mathfrak{N}_1 \ldots \lhd \mathfrak{N}_s = \mathfrak{G}$ eine Hauptreihe von $\mathfrak{G}$, so bewirkt jeder innere Automorphismus $X \to GXG^{-1}$ von $\mathfrak{G}$ einen Automorphismus auf den Hauptfaktoren $\mathfrak{N}_j/\mathfrak{N}_{j-1}$. Diese Hauptfaktoren sind bekanntlich elementar abelsche Gruppen. Legen wir für $\mathfrak{N}_j/\mathfrak{N}_{j-1}$ eine Basis $T_1, T_2, \ldots, T_{n_j}$ zugrunde, so wird ein Automorphismus in seiner Wirkung auf die Basiselemente T_i beschrieben durch

$$T_i \to T_1^{\xi_{i1}} T_2^{\xi_{i2}} \ldots T_{n_j}^{\xi_{in_j}}.$$

Gehen wir von der multiplikativen Schreibweise der abelschen Gruppe $\mathfrak{N}_j/\mathfrak{N}_{j-1}$ zur additiven Schreibweise über, so wird der Automorphismus zu einer linearen Abbildung eines Moduls. Hat $\mathfrak{N}_j/\mathfrak{N}_{j-1}$ die Ordnung $p_j^{n_j}$, so ist dieser Modul der n_j-dimensionale Vektorraum über dem Galois-Feld $GF(p_j)$. Damit haben wir eine Beschreibung der Wirkung der inneren Automorphismen von $\mathfrak{G}$ auf die Hauptfaktoren von $\mathfrak{G}$ durch lineare Gruppen erhalten. Zu beachten ist, daß es sich dabei um Gruppen über endlichen Körpern handelt.

Für eine genauere Kenntnis der Struktur abstrakter auflösbarer Gruppen ist also die Untersuchung der Eigenschaften von linearen auflösbaren Gruppen unerläßlich. Insbesondere ist es von Wichtigkeit, gewisse Strukturinvarianten einer linearen auflösbaren Gruppe durch den Grad n abzuschätzen. Die wichtigsten Invarianten sind folgende:

Die Stufe $k(\mathfrak{G})$; $k(\mathfrak{G})$ ist die kleinste ganze Zahl k mit $\mathfrak{G}^{(k)} = \mathfrak{E}$ ($\mathfrak{G}^{(k)} = k$-malige Bildung der Kommutatorgruppe).

Der Rang $r(\mathfrak{G})$; haben die Hauptfaktoren $\mathfrak{N}_j/\mathfrak{N}_{j-1}$ die Ordnungen $p_j^{n_j}$, so ist $r(\mathfrak{G}) = \max_j n_j$.

Die höchste Potenz p^g der Charakteristik p, welche in der Ordnung einer über $GF(p^l)$ vollständig reduziblen, auflösbaren Gruppe aufgeht.

[1] Diese Arbeit ist der erste Teil meiner Habilitationsschrift (Universität Tübingen); der zweite Teil wird demnächst unter dem Titel „Zweifach transitive, auflösbare Permutationsgruppen" erscheinen.

Nach vorbereitenden Bemerkungen in § 1 über das Zerfallen von Normalteilern irreduzibler linearer Gruppen beschreiben wir in § 2 das in den meisten späteren Beweisen verwendete Induktionsprinzip, welches eine Induktion nach dem Grad n der Gruppe gestattet. Dieses Prinzip wurde teilweise bereits von C. JORDAN verwandt, um ein Konstruktionsverfahren für lineare auflösbare Gruppen anzugeben (C. JORDAN [12]). Wir benutzen es umgekehrt zur Strukturanalyse von linearen auflösbaren Gruppen. In § 3 folgen einfache Anwendungen auf die Darstellungen auflösbarer Gruppen (monomiale Darstellungen, Grade primitiver Darstellungen). In § 4 beginnt dann mit der Abschätzung der Stufe $k(\mathfrak{G})$ der Kern der Untersuchung. Es ergibt sich $k(\mathfrak{G}) \leq \text{Min}\{2n, 7\,{}_2\log n + 1\}$. In § 5 folgt für vollständig reduzible Gruppen der Nachweis von $r(\mathfrak{G}) \leq n$, darauf in § 6 für vollständig reduzible Gruppen über $GF(p^f)$ der Beweis von $g \leq f(3n \to p)/p$. Der letztere Beweis ist weit schwieriger als die vorhergehenden; denn nun hängt die Aussage wesentlich vom zugrunde liegenden Körper $GF(p^f)$ ab, was feinere Untersuchungen arithmetischer Natur erfordert.

In § 7 werden die erhaltenen Resultate auf abstrakte auflösbare Gruppen angewandt. Folgendes sind die Hauptergebnisse:

Ist $r(\mathfrak{G}) = r$, so ist $\mathfrak{G}^{(2^r)}$ nilpotent. Dies ist ein Analogon zu dem bekannten Satz von E. WENDT [19], daß für überauflösbare Gruppen [d.h. für $r(\mathfrak{G}) = 1$] die Kommutatorgruppe $\mathfrak{G}'$ nilpotent ist. Ferner ergibt sich eine Beziehung zwischen den wichtigen Invarianten $r_p(\mathfrak{G})$ (p-Rang) und $l_p(\mathfrak{G})$ (p-Länge), welche in folgender Weise definiert sind:

$r_p(\mathfrak{G})$: Ist $\mathfrak{E} = \mathfrak{N}_0 \lhd \mathfrak{N}_1 \lhd \cdots \lhd \mathfrak{N}_s = \mathfrak{G}$ eine Hauptreihe von $\mathfrak{G}$ mit Hauptfaktoren $\mathfrak{N}_j/\mathfrak{N}_{j-1}$ der Ordnungen $p_j^{n_j}$, so sei $r_p(\mathfrak{G}) = \underset{p_j = p}{\text{Max}}\, n_j$. [Offenbar gilt $r(\mathfrak{G}) = \underset{p}{\text{Max}}\, r_p(\mathfrak{G})$.]

$l_p(\mathfrak{G})$: Sei p ein fester Primteiler der Ordnung von $\mathfrak{G}$ und $\mathfrak{E} = \mathfrak{M}_0 \lhd \mathfrak{M}_1 \lhd \cdots \lhd \mathfrak{M}_t = \mathfrak{G}$ eine Normalteilerkette von $\mathfrak{G}$ mit der Eigenschaft, daß die Ordnungen der Faktorgruppen $\mathfrak{M}_j/\mathfrak{M}_{j-1}$ entweder zu p teilerfremd sind oder Potenzen von p sind. In dieser Kette mögen genau s Faktorgruppen von p-Potenzordnung auftreten. Dann bezeichnen wir das kleinste s, welches bei irgendeiner solchen Kette vorkommt, als p-Länge $l_p(\mathfrak{G})$.

Mit Hilfe der oben erwähnten Abschätzung von g ergibt sich eine Abschätzung von $l_p(\mathfrak{G})$ durch $r_p(\mathfrak{G})$, nämlich $l_p(\mathfrak{G}) \leq r_p(\mathfrak{G})$ für alle p, für $p > 3$ sogar

$$l_p(\mathfrak{G}) \leq \left[{}_{p/3}\log \frac{r_p(\mathfrak{G})}{p-1} \right] + 2.$$

Dieses Resultat ist besonders interessant, da P. HALL und G. HIGMAN in [8] kürzlich $l_p(\mathfrak{G})$ für p-auflösbare Gruppen durch gewisse Invarianten der p-Sylow-Gruppen (Klasse, Stufe, Exponent) abgeschätzt haben. Dabei trat die Bedeutung von $l_p(\mathfrak{G})$ zum ersten Male ans Licht.

Herrn H. WIELANDT danke ich herzlich für zahlreiche Ratschläge bei der Niederschrift dieser Arbeit.

Bezeichnungen

$|\mathfrak{G}|$ = Ordnung von $\mathfrak{G}$; $\mathfrak{U} \leq \mathfrak{G}$ bedeutet: $\mathfrak{U}$ ist Untergruppe von $\mathfrak{G}$, wobei auch $\mathfrak{U} = \mathfrak{G}$ zugelassen ist; $\mathfrak{U} < \mathfrak{G}$: $\mathfrak{U}$ ist Untergruppe von $\mathfrak{G}$, aber $\mathfrak{U} \neq \mathfrak{G}$; $\mathfrak{U} \trianglelefteq \mathfrak{G}$: $\mathfrak{U}$ ist Normalteiler von $\mathfrak{G}$; $\mathfrak{U} \triangleleft \mathfrak{G}$: $\mathfrak{U}$ ist Normalteiler von $\mathfrak{G}$ und $\mathfrak{U} \neq \mathfrak{G}$; $A(\mathfrak{G})$ = Automorphismengruppe von $\mathfrak{G}$; $\mathfrak{G}_p$ = p-Sylow-Gruppe von $\mathfrak{G}$; ${}_p\mathfrak{G}$ = p-Sylow-Komplement von $\mathfrak{G}$; $\mathfrak{Z}(\mathfrak{G})$ = Zentrum von $\mathfrak{G}$; $\mathfrak{G}'$ = Kommutatorgruppe von $\mathfrak{G}$; $\mathfrak{G}^{(n)}$ = n-te Kommutatorgruppe von $\mathfrak{G}$; $(X, Y) = XYX^{-1}Y^{-1}$; $\Phi(\mathfrak{G})$ = Frattini-Gruppe von $\mathfrak{G}$; $\mathfrak{F}$ = maximaler nilpotenter Normalteiler von $\mathfrak{G}$; $\mathfrak{F}(p)$ = maximaler p-nilpotenter Normalteiler von $\mathfrak{G}$; $\mathfrak{N}(\mathfrak{U})$ = Normalisator der Untergruppe $\mathfrak{U}$ von $\mathfrak{G}$; $\mathfrak{C}(\mathfrak{U})$ = Zentralisator der Untergruppe $\mathfrak{U}$ von $\mathfrak{G}$; $r(\mathfrak{G})$ = Rang von $\mathfrak{G}$ (s. Einleitung); $r_{\mathfrak{G}}(\mathfrak{N})$ = Relativrang von $\mathfrak{N}$ in $\mathfrak{G}$: Ist $\mathfrak{E} = \mathfrak{N}_0 \triangleleft \mathfrak{N}_1 \triangleleft \cdots \triangleleft \mathfrak{N}_s = \mathfrak{N}$ das Stück einer Hauptreihe von $\mathfrak{G}$ unterhalb von $\mathfrak{N}$ und ist $|\mathfrak{N}_j/\mathfrak{N}_{j-1}| = p_j^{a_j}$, so sei $r_{\mathfrak{G}}(\mathfrak{N}) = \max_j a_j$; $r_p(\mathfrak{G})$ = p-Rang von $\mathfrak{G}$ (s. Einleitung); $l_p(\mathfrak{G})$ = p-Länge von $\mathfrak{G}$ (s. Einleitung); mit p und q bezeichnen wir stets Primzahlen; ist $p^a \mid m$, aber $p^{a+1} \nmid m$, so schreiben wir $p^a \top m$; $[x]$ sei die größte ganze Zahl k mit $k \leq x$.

§ 1. Reduzible Normalteiler irreduzibler Gruppen

CLIFFORD hat in [3] allgemein untersucht, in welcher Weise ein Normalteiler $\mathfrak{N}$ einer irreduziblen Gruppe $\mathfrak{G}$ reduzibel sein kann. Wir stellen in diesem Paragraphen die wichtigsten Ergebnisse mit kurzen Beweisen zusammen und ziehen einige Folgerungen daraus. Dabei werden wir die Endlichkeit von $|\mathfrak{G}|$ nicht verwenden, nur an einigen Stellen die Endlichkeit von $|\mathfrak{G}/\mathfrak{N}|$. Der Körper K ist beliebig, sofern nicht ausdrücklich eine Voraussetzung über K angegeben wird.

Satz 1. Ist $\mathfrak{G}$ irreduzibel über K, so zerfällt jeder reduzible Normalteiler $\mathfrak{N}$ von $\mathfrak{G}$ über K vollständig in irreduzible Bestandteile des gleichen Grades.

Beweis. Sei $\mathfrak{m}$ der Darstellungsmodul von $\mathfrak{G}$, $\mathfrak{n}$ ein bezüglich $\mathfrak{N}$ invarianter und irreduzibler Teilmodul. Ist G ein beliebiges Element von $\mathfrak{G}$, so ist auch $G\mathfrak{n}$ unter $\mathfrak{N}$ invariant; denn wegen $G^{-1}NG \in \mathfrak{N}$ haben wir $NG\mathfrak{n} = GG^{-1}NG\mathfrak{n} = G\mathfrak{n}$. Aus der Irreduzibilität von $\mathfrak{N}$ auf $\mathfrak{n}$ folgt die Irreduzibilität von $G\mathfrak{n}$ bezüglich $\mathfrak{N}$; denn ein unter $\mathfrak{N}$ invarianter Teilmodul von $G\mathfrak{n}$ ginge bei Anwendung von G^{-1} in einen solchen von $\mathfrak{n}$ über. Da $\mathfrak{G}$ auf $\mathfrak{m}$ irreduzibel ist, spannen die $G\mathfrak{n}$ ganz $\mathfrak{m}$ auf. Berücksichtigen wir, daß die verschiedenen $G\mathfrak{n}$ paarweise den Durchschnitt 0 haben, so ist nun leicht zu sehen, daß der ganze Darstellungsmodul $\mathfrak{m}$ die direkte Summe geeigneter $G\mathfrak{n}$ ist, also $\mathfrak{N}$ vollständig in irreduzible Teile gleichen Grades zerfällt.

Der Beweis zeigt noch, daß die Anzahl der Bestandteile von $\mathfrak{N}$ nicht größer sein kann als der Index von $\mathfrak{N}$ in $\mathfrak{G}$.

Satz 2. $\mathfrak{G}$ sei irreduzibel über K, $\mathfrak{N} \triangleleft \mathfrak{G}$; $\mathfrak{N}$ zerfalle über K in irreduzible Bestandteile, die nicht alle äquivalent sind. Fassen wir die Teilmoduln $\mathfrak{n}_i$ des Darstellungsmoduls $\mathfrak{m}$, auf welchen $\mathfrak{N}$ äquivalente, über K irreduzible

Darstellungen bewirkt, zu Teilmoduln $\mathfrak{r}_i$ zusammen, so werden die $\mathfrak{r}_i$ von $\mathfrak{G}$ transitiv vertauscht, sind also Imprimitivitätsgebiete von $\mathfrak{G}$ (SPEISER [15], S. 197). Ist $\mathfrak{G}_1$ diejenige Untergruppe von $\mathfrak{G}$, welche $\mathfrak{r}_1$ in sich überführt, so hat $\mathfrak{G}_1$ die Gestalt

$$\mathfrak{G}_1 = \begin{pmatrix} B(\mathfrak{G}_1) & 0 & \dots & 0 \\ 0 & & & \\ \vdots & & * & \\ 0 & & & \end{pmatrix}$$

(in geeigneten Koordinaten), und $B(\mathfrak{G}_1)$ ist irreduzibel über K.

Beweis. $\mathfrak{N}$ besitze die über K irreduziblen, inäquivalenten Bestandteile $F_i(\mathfrak{N})$ jeweils mit der Vielfachheit j_i. Der Teilmodul $\mathfrak{r}_1$ von $\mathfrak{m}$ erfahre die Darstellung $F_1(\mathfrak{N})$ mit der Vielfachheit j_1. Aus $GN = GNG^{-1}G = N^{G^{-1}}G$ folgt $GN\mathfrak{r}_1 = G\mathfrak{r}_1 = N^{G^{-1}}(G\mathfrak{r}_1)$. Somit ist $G\mathfrak{r}_1$ invariant unter $\mathfrak{N}$, und die Operatormoduln $(\mathfrak{r}_1|\mathfrak{N})$ und $(G\mathfrak{r}_1|\mathfrak{N}^{G^{-1}})$ sind isomorph. Daher zerfällt auch $G\mathfrak{r}_1$ unter $\mathfrak{N}$ in j_1 äquivalente irreduzible Teile. Wegen der Irreduzibilität von $\mathfrak{G}$ spannen die $G\mathfrak{r}_i$ ganz $\mathfrak{m}$ auf.

Ist $\mathfrak{d} = G\mathfrak{r}_1 \cap H\mathfrak{r}_1 \neq 0$, so hat $\mathfrak{N}$ auf $\mathfrak{d}$, $G\mathfrak{r}_1$ und $H\mathfrak{r}_1$ nur äquivalente irreduzible Bestandteile $F_i(\mathfrak{N})$. Da $G\mathfrak{r}_1$ [wie $\mathfrak{r}_1$ für $F_1(\mathfrak{N})$] ein maximaler Teilmodul mit dieser Eigenschaft ist, folgt $G\mathfrak{r}_1 = H\mathfrak{r}_1$.

Nun folgt leicht, daß $\mathfrak{m}$ die direkte Summe derjenigen $\mathfrak{r}_i = G\mathfrak{r}_1$ ist, welche Vielfache verschiedener irreduzibler Darstellungen $F_i(\mathfrak{N})$ unter $\mathfrak{N}$ erfahren. Diese $\mathfrak{r}_i$ werden von $\mathfrak{G}$ vertauscht, und zwar transitiv, da $\mathfrak{G}$ über K irreduzibel ist.

Nun zeigen wir noch, daß $\mathfrak{r}_1$ unter $B(\mathfrak{G}_1)$ irreduzibel ist. Sei $\mathfrak{s}$ ein unter $B(\mathfrak{G}_1)$ invarianter, von 0 verschiedener Teilmodul von $\mathfrak{r}_1$. Sind $G_1 = E$, $G_2, \dots, G_k$ Vertreter von $\mathfrak{G}_1$ in $\mathfrak{G}$, so ist die direkte Summe $G_1\mathfrak{s} + \cdots + G_k\mathfrak{s}$ der $G_i\mathfrak{s}$ invariant unter $\mathfrak{G}$, fällt daher mit $\mathfrak{m}$ zusammen. Aus $\mathfrak{m} = G_1\mathfrak{r}_1 + \cdots + G_k\mathfrak{r}_1$ und $G_i\mathfrak{s} \subseteq G_i\mathfrak{r}_1$ folgt nun $G_1\mathfrak{r}_1 = \mathfrak{r}_1 = G_1\mathfrak{s} = \mathfrak{s}$, also die Irreduzibilität von $B(\mathfrak{G}_1)$ auf $\mathfrak{r}_1$.

Während die bisherigen Resultate nur die Irreduzibilität von $\mathfrak{G}$ über einem beliebigen Körper K erfordern, benötigen wir im Beweis von Satz 3 das Schursche Lemma mit seiner scharfen Aussage für irreduzible Gruppen über algebraisch abgeschlossenen Körpern, müssen daher die Irreduzibilität der Bestandteile des Normalteilers $\mathfrak{N}$ beim Übergang zur algebraisch abgeschlossenen Hülle von K voraussetzen.

Satz 3. $\mathfrak{G}$ sei irreduzibel über K; der Normalteiler $\mathfrak{N}$ von $\mathfrak{G}$ zerfalle in n_1 äquivalente Teile $A(\mathfrak{N})$ vom Grad n_2, welche noch über der algebraisch abgeschlossenen Hülle von K irreduzibel sind. Dann läßt jedes Element $G \in \mathfrak{G}$ sich zerlegen in ein Kronecker-Produkt $G = G_1 \times G_2$ mit Grad $G_i = n_i$. Die G_i sind bis auf Faktoren aus K bestimmt und bilden projektive, über K irreduzible Darstellungen $P_i(\mathfrak{G})$ von $\mathfrak{G}$. P_1 ist sogar eine projektive Darstellung von $\mathfrak{G}/\mathfrak{N}$.

Beweis. Bei geeignetem Koordinatensystem ist

$$\mathfrak{N} = \begin{pmatrix} A(\mathfrak{N}) & 0 & \dots & 0 \\ 0 & A(\mathfrak{N}) & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & A(\mathfrak{N}) \end{pmatrix}$$

[n_1 Bestandteile $A(\mathfrak{N})$]. Mit der entsprechenden Unterteilung $G = (G_{ij})$ $(i, j = 1, \dots, n_1)$ wird dann

$$G_{ij} A(N) = A(GNG^{-1}) G_{ij}.$$

Da nicht alle G_{ij} verschwinden, sind $A(\mathfrak{N})$ und $A(G\mathfrak{N}G^{-1})$ nach dem Schurschen Lemma äquivalent, und das Element G_{ij} ist bis auf einen skalaren Faktor bestimmt, also $G_{ij} = \mathfrak{p}_{ij} P$ und $G = (\mathfrak{p}_{ij}) \times P$. Die Matrizen $(\mathfrak{p}_{ij})$ und P bilden, wie leicht zu sehen ist, projektive Darstellungen $P_1(\mathfrak{G})$ und $P_2(\mathfrak{G})$ von $\mathfrak{G}$ der Grade n_1 und n_2. Die Gestalt von $\mathfrak{N}$ zeigt, daß auf $\mathfrak{N}$ gilt $(\mathfrak{p}_{ij}) = c E_{n_1}$, also P_1 eine Darstellung von $\mathfrak{G}/\mathfrak{N}$ ist (mit E_{n_1} bezeichnen wir dabei die n_1-reihige Einheitsmatrix). Die Irreduzibilität der $P_i(\mathfrak{G})$ über K folgt sofort aus der Irreduzibilität von $\mathfrak{G}$ über K.

Die Sätze 1 bis 3 werden im folgenden immer wieder herangezogen werden. Bevor wir zu einfachen Anwendungen übergehen, geben wir noch drei bekannte Hilfssätze an:

Hilfssatz 1. Sei $\mathfrak{G}$ irreduzibel über dem Körper K der Charakteristik p. Dann besitzt $\mathfrak{G}$ keinen Normalteiler von einer p-Potenzordnung $p^a > 1$.

Beweis. Angenommen, es sei $\mathfrak{N} \trianglelefteq \mathfrak{G}$ und $|\mathfrak{N}| = p^a > 1$. Dann ist auch das Zentrum $\mathfrak{Z}(\mathfrak{N})$ von $\mathfrak{N}$ ein Normalteiler von $\mathfrak{G}$ mit $|\mathfrak{Z}(\mathfrak{N})| = p^b > 1$. Die absolut irreduziblen Darstellungen von $\mathfrak{Z}(\mathfrak{N})$ haben den Grad 1. Da kein Erweiterungskörper von K primitive p-te Einheitswurzeln enthält, ist die identische Darstellung die einzige absolut irreduzible Darstellung von $\mathfrak{Z}(\mathfrak{N})$, und diese läßt sich schon über K angeben. Also läßt sich $\mathfrak{Z}(\mathfrak{N})$ über K auf Dreiecksgestalt bringen. Der Teilmodul $\mathfrak{n}$ des Darstellungsmoduls $\mathfrak{m}$, auf welchem $\mathfrak{Z}(\mathfrak{N})$ die identische Darstellung bewirkt, erfüllt daher $0 < \mathfrak{n} < \mathfrak{m}$ und ist invariant unter $\mathfrak{G}$, entgegen der Irreduzibilität von $\mathfrak{G}$. Damit haben wir den gewünschten Widerspruch erhalten.

Hilfssatz 2. Sei $\mathfrak{A}$ eine abelsche, über $GF(p^l)$ irreduzible Gruppe. Dann ist $\mathfrak{A}$ zyklisch. Ist A ein erzeugendes Element von $\mathfrak{A}$, so ist das charakteristische Polynom $\varphi(x)$ von A irreduzibel über $GF(p^l)$.

Beweis. Siehe HUPPERT [9], S. 416, Hilfssatz 4. Dort wird der Satz nur für $GF(p)$ formuliert, aber der Beweis überträgt sich unverändert auf den vorliegenden allgemeineren Fall.

Hilfssatz 3. Sei $\mathfrak{G}$ irreduzibel über dem Körper K. Das Zentrum der von den Matrizen von $\mathfrak{G}$ erzeugten einfachen Algebra sei separabel über K. Ist L eine galoissche Erweiterung von K, so haben die irreduziblen Bestandteile von $\mathfrak{G}$ über L alle den gleichen Grad und sind treu. Wählen wir L

genügend groß, so können wir erreichen, daß $\mathfrak{G}$ über L in absolut irreduzible, treue Bestandteile gleichen Grades vollständig zerfällt.

Beweis. Unter den vorliegenden Umständen zerfällt $\mathfrak{G}$ vollständig über L in solche Bestandteile, die durch Automorphismen von L über K ineinander übergehen, insbesondere also gleichen Grad haben und treu sind [v. d. WAERDEN [17], S. 69 (3)]. Wählen wir L als galoisschen Zerfällungskörper der zu $\mathfrak{G}$ gehörigen Algebra, so folgt die zweite Behauptung.

Bei den späteren Anwendungen von Hilfssatz 3 wird K entweder der komplexe Zahlkörper oder ein Galois-Feld sein. Die Voraussetzung, daß das Zentrum der von $\mathfrak{G}$ erzeugten Algebra über K separabel ist, ist dann sicher erfüllt.

Wir kommen zu ersten Anwendungen:

Satz 4. $\mathfrak{G}$ sei irreduzibel über dem algebraisch abgeschlossenen Körper K der Charakteristik 0; der Normalteiler $\mathfrak{N}$ von $\mathfrak{G}$ zerfalle in n absolut irreduzible Teile. Ist $|\mathfrak{G}/\mathfrak{N}|$ endlich, so ist n ein Teiler von $|\mathfrak{G}/\mathfrak{N}|$.

Beweis. Entsprechend den Sätzen 1 und 2 sei

$$\mathfrak{N} = \begin{pmatrix} E_k \times A_1(\mathfrak{N}) & 0 & \dots & 0 \\ 0 & E_k \times A_2(\mathfrak{N}) & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & E_k \times A_m(\mathfrak{N}) \end{pmatrix}$$

mit $n = km$ und inäquivalenten $A_i(\mathfrak{N})$. Ist $m > 1$, so ist $\mathfrak{G}$ nach Satz 2 imprimitiv. Die Teilmoduln $\mathfrak{r}_i$ $(i = 1, \dots, m)$ werden von $\mathfrak{G}$ transitiv permutiert. Ist $\mathfrak{G}_1$ diejenige Untergruppe von $\mathfrak{G}$, welche $\mathfrak{r}_1$ fest läßt, so hat $\mathfrak{G}_1$ die Gestalt

$$\mathfrak{G}_1 = \begin{pmatrix} B(\mathfrak{G}_1) & 0 & \dots & 0 \\ 0 & & & \\ \vdots & & * & \\ 0 & & & \end{pmatrix},$$

und nach einem bekannten Satz über transitive Permutationsgruppen ist $m = |\mathfrak{G} : \mathfrak{G}_1|$.

Nun betrachten wir die Darstellung $B(\mathfrak{G}_1)$ von $\mathfrak{G}_1$, unter der $\mathfrak{N}$ in k Teile $A_1(\mathfrak{N})$ zerfällt. $B(\mathfrak{G}_1)$ ist nach Satz 2 irreduzibel über dem algebraisch abgeschlossenen Körper K. Für $B(\mathfrak{G}_1)$ liegt also die Situation von Satz 3 vor, woraus $B(\mathfrak{G}_1) = P_1(\mathfrak{G}_1) \times P_2(\mathfrak{G}_1)$ mit projektiven, absolut irreduziblen Darstellungen $P_i(\mathfrak{G}_1)$ folgt. $P_1(\mathfrak{G}_1)$ ist nach Satz 3 sogar eine projektive Darstellung der endlichen Gruppe $\mathfrak{G}_1/\mathfrak{N}$ vom Grad k und ist irreduzibel über K. Nach SCHUR [14] (Satz VII a) ist dann $k \mid |\mathfrak{G}_1/\mathfrak{N}|$. Insgesamt folgt jetzt

$$n = km \mid |\mathfrak{G}_1/\mathfrak{N}| \, |\mathfrak{G} : \mathfrak{G}_1| = |\mathfrak{G}/\mathfrak{N}|,$$

also die Behauptung.

Für den Spezialfall, daß $\mathfrak{G}/\mathfrak{N}$ zyklisch ist, findet sich dieser Satz auch in CLIFFORD [3].

Aus Satz 4 folgt nun leicht ein von N. ITÔ für endliche Gruppen angegebener Satz über die Grade absolut irreduzibler Darstellungen (N. ITÔ [10]):

Satz 5. Sei $\mathfrak{G}$ irreduzibel vom Grad n über dem algebraisch abgeschlossenen Körper K der Charakteristik 0, $\mathfrak{A}$ ein abelscher Normalteiler von $\mathfrak{G}$. Ist $|\mathfrak{G}/\mathfrak{A}|$ endlich, so gilt $n \mid |\mathfrak{G}/\mathfrak{A}|$.

Beweis. Da $\mathfrak{A}$ abelsch ist, zerfällt es in n Teile vom Grad 1. Nach Satz 4 folgt nun $n \mid |\mathfrak{G}/\mathfrak{A}|$.

N. ITÔ hatte seinen Satz durch eine geschickte Rechnung mit Charakteren bewiesen. In unseren Beweis scheinen die Charaktere nicht einzugehen; aber die Herleitung des im Beweis von Satz 4 verwendeten Satzes von SCHUR benötigt Gruppencharaktere, so daß unser Beweis von Satz 5 nicht „charakterfrei" ist.

§ 2. Das Induktionsprinzip

Bei einem großen Teil der Beweise in den folgenden Paragraphen verwenden wir einen Induktionsschluß nach dem Grad n der auflösbaren Gruppe $\mathfrak{G}$. Für Gruppen vom Grad 1 ist die Behauptung jeweils trivial, so daß die Induktionsbasis stets gesichert ist. Der Grundgedanke des Induktionsverfahrens ist folgender: Die Sätze 1 bis 3 liefern Informationen über $\mathfrak{G}$, sofern reduzible Normalteiler von $\mathfrak{G}$ vorhanden sind. Da eine auflösbare Gruppe $\mathfrak{G}$ stets nichttriviale abelsche Normalteiler besitzt, welche über geeigneten Körpern reduzibel sind, können wir hoffen, auf diese Weise die auflösbaren linearen Gruppen in die Hand zu bekommen. Bei der Beschreibung des Induktionsverfahrens beschränken wir uns der Einfachheit halber auf endliche Gruppen $\mathfrak{G}$, die über einem algebraisch abgeschlossenen Körper K irreduzibel sind. Jedoch formulieren wir die grundlegenden Hilfssätze I bis III gleich so, daß sie auch die späteren Fälle mit umfassen.

Je nach dem Zerfallen der Normalteiler von $\mathfrak{G}$ gehen wir verschieden vor:

Fall 1. $\mathfrak{G}$ enthalte einen Normalteiler $\mathfrak{N}$, der inäquivalente irreduzible Bestandteile besitzt. Nach Satz 2 ist dann $\mathfrak{G}$ imprimitiv. Sei $\mathfrak{D}$ derjenige Normalteiler von $\mathfrak{G}$, welcher alle Imprimitivitätsgebiete $\mathfrak{r}_i$ fest läßt. Ist k die Anzahl der Imprimitivitätsgebiete, so zerfällt $\mathfrak{D}$ vollständig in k Teile vom Grad n/k. Auf die Bestandteile von $\mathfrak{D}$ können wir nun unsere jeweilige Induktionsannahme anwenden. Die Faktorgruppe $\mathfrak{G}/\mathfrak{D}$ ist isomorph zu einer Permutationsgruppe $\mathfrak{P}_k$ vom Grad k, welche wir als lineare Gruppe vom Grad k über dem komplexen Zahlkörper C auffassen wollen. Ist $k<n$, so haben wir in $\mathfrak{P}_k$ eine Gruppe von kleinerem Grad erhalten. Für $k=n$ (d.h. $\mathfrak{G}$ monomial) ist noch eine besondere Betrachtung nötig: $\mathfrak{P}_k$ ist vollständig reduzibel und enthält mindestens einmal die identische Darstellung, ist daher isomorph zu einer Gruppe vom Grad $k-1$. Damit sind wir wieder zu einer Gruppe kleineren Grades gekommen, auf welche wir die Induktionsannahme anwenden können.

Schließlich sind noch die Resultate für $\mathfrak{D}$ und $\mathfrak{G}/\mathfrak{D} \cong \mathfrak{P}_k$ zu der gewünschten Aussage über $\mathfrak{G}$ zusammenzusetzen.

Fall 2. Nun habe $\mathfrak{G}$ einen Normalteiler $\mathfrak{N}$, der in $n_1 > 1$ absolut irreduzible gleiche Teile vom Grad $n_2 > 1$ zerfällt. Nach Satz 3 läßt sich dann $\mathfrak{G}$ zerlegen in ein Kronecker-Produkt $\mathfrak{G} = P_1(\mathfrak{G}) \times P_2(\mathfrak{G})$. Indem wir durch Übergang zu einer assoziierten Darstellung von $P_i(\mathfrak{G})$ die Darstellungsmatrizen auf Determinante 1 normieren, erreichen wir, daß das Faktorensystem $a_{S,T}$ aus lauter n-ten Einheitswurzeln besteht[2]). Nun nehmen wir zu den Darstellungsmatrizen von $P_i(\mathfrak{G})$ alle Matrizen $a_{S,T}E$ hinzu. So entsteht eine lineare Gruppe $\mathfrak{K}_i = \mathfrak{K}(P_i(\mathfrak{G}))$. Dabei ist $\mathfrak{K}_i$ eine zentrale Erweiterung einer endlichen zyklischen Gruppe mit einem homomorphen Bild von $\mathfrak{G}$, ist daher endlich und auflösbar[3]). Da die $\mathfrak{K}_i$ die Grade $n_i < n$ haben, können wir unsere Induktionsannahme auf sie anwenden. Es bleibt dann die Aufgabe, aus den Resultaten über die $\mathfrak{K}_i$ die gewünschte Aussage für $\mathfrak{G}$ zu folgern.

Fall 3. Schließlich ist es noch möglich, daß jeder reduzible Normalteiler von $\mathfrak{G}$ lauter gleiche Bestandteile vom Grad 1 hat, also im Zentrum $\mathfrak{Z}(\mathfrak{G})$ von $\mathfrak{G}$ liegt. Insbesondere liegt dann jeder abelsche Normalteiler von $\mathfrak{G}$ in $\mathfrak{Z}(\mathfrak{G})$. Dieser Fall, in dem die reduziblen Normalteiler von $\mathfrak{G}$ keine nützliche Aussage über $\mathfrak{G}$ liefern, erfordert eine eingehende Betrachtung der minimalen nichtabelschen Normalteiler von $\mathfrak{G}$.

Die nun vorliegenden Verhältnisse beschreiben wir in drei wichtigen Hilfssätzen, welche für Anwendungen in einer späteren Arbeit allgemeiner formuliert werden, als es im Augenblick nötig ist; bei den Anwendungen in der vorliegenden Arbeit wird die Gruppe $\mathfrak{A}$ stets aus Matrizen der Gestalt aE bestehen, so daß dann $\mathfrak{Z}(\mathfrak{G}) = \mathfrak{A}$ gilt.

Hilfssatz I (Struktur minimaler nichtabelscher Normalteiler).

Voraussetzung. Sei $\mathfrak{G}$ auflösbar, nichtabelsch, linear vom Grad n über dem Körper K. Sei $\mathfrak{A}$ ein maximaler abelscher Normalteiler von $\mathfrak{G}$; $\mathfrak{A}$ sei zyklisch.

Behauptung. Ist $\mathfrak{N}$ ein endlicher, minimaler nichtabelscher Normalteiler von $\mathfrak{G}$ und $\mathfrak{N} \leq \mathfrak{C}(\mathfrak{A})$, so gilt

1. $\mathfrak{N}$ ist nilpotent von der Klasse 2 und hat Primzahlpotenzordnung q^b.
2. $\mathfrak{N}/\mathfrak{Z}(\mathfrak{N})$ ist ein minimaler Normalteiler von $\mathfrak{G}/\mathfrak{Z}(\mathfrak{N})$, ist daher elementar abelsch.
3. $\mathfrak{N}'$ hat die Ordnung q und liegt, ebenso wie $\mathfrak{Z}(\mathfrak{N})$, in $\mathfrak{A}$.
4. Ist $q > 2$, so gilt für alle $N \in \mathfrak{N}$ die Gleichung $N^q = E$.
5. Ist $q = 2$, so gilt für $N \in \mathfrak{N}$ stets $N^4 = E$.
6. Ist $\mathfrak{G}$ irreduzibel über K, so gilt Char. $K \neq q$. Ist $K = GF(p^l)$ und besteht $\mathfrak{A}$ nur aus Matrizen der Gestalt aE, so ist $q \mid p^l - 1$.

---

[2]) Die $a_{S,T}$ sind erklärt durch $P_i(S)\,P_i(T) = a_{S,T}^{(i)} P_i(ST)$ (SCHUR [14]).

[3]) Die Konstruktion von $\mathfrak{K}(P_i(\mathfrak{G}))$ ist der Schurschen Konstruktion der Darstellungsgruppe nachgebildet, führt aber in dieser einfachen Form nur zu einer von P_i abhängigen Gruppe $\mathfrak{K}_i$ (SCHUR [14]).

Beweis.

1. Die Kommutatorgruppe $\mathfrak{N}'$ von $\mathfrak{N}$ ist von $\mathfrak{N}$ verschieden und invariant in $\mathfrak{G}$. Wegen der Wahl von $\mathfrak{N}$ ist daher $\mathfrak{N}'$ abelsch. Aus $\mathfrak{N}' \leq \mathfrak{N} \leq \mathfrak{C}(\mathfrak{A})$ folgt nun, daß auch $\mathfrak{A}\mathfrak{N}'$ ein abelscher Normalteiler von $\mathfrak{G}$ ist. Da aber $\mathfrak{A}$ ein maximaler abelscher Normalteiler ist, erhalten wir $\mathfrak{N}' \leq \mathfrak{A}$. Insbesondere ist dann $\mathfrak{N}' \leq \mathfrak{Z}(\mathfrak{N})$, also $\mathfrak{N}$ nilpotent von der Klasse 2. $\mathfrak{N}$ besitzt mindestens eine nichtabelsche Sylow-Gruppe $\mathfrak{N}_q$. Diese ist invariant in $\mathfrak{G}$, fällt daher mit $\mathfrak{N}$ zusammen. Damit haben wir $\mathfrak{N} = \mathfrak{N}_q$ und $|\mathfrak{N}| = q^b$.

2. Ist $\mathfrak{U}$ ein Normalteiler von $\mathfrak{G}$ mit $\mathfrak{U} < \mathfrak{N}$, so ist $\mathfrak{U}$ wegen der Wahl von $\mathfrak{N}$ abelsch. Da $\mathfrak{U}$ mit $\mathfrak{N}$ in $\mathfrak{C}(\mathfrak{A})$ liegt, ist wieder $\mathfrak{A}\mathfrak{U}$ abelsch, also wie unter 1. $\mathfrak{U} \leq \mathfrak{A}$. Es folgt $\mathfrak{U} \leq \mathfrak{N} \cap \mathfrak{A} \leq \mathfrak{Z}(\mathfrak{N})$, und insbesondere für $\mathfrak{U} = \mathfrak{Z}(\mathfrak{N})$ dann $\mathfrak{Z}(\mathfrak{N}) \leq \mathfrak{A}$. Also ist $\mathfrak{N}/\mathfrak{Z}(\mathfrak{N})$ ein minimaler Normalteiler von $\mathfrak{G}/\mathfrak{Z}(\mathfrak{N})$, daher elementar abelsch. Für jedes $N \in \mathfrak{N}$ gilt dann $N^q \in \mathfrak{Z}(\mathfrak{N})$.

3. Die Behauptung $\mathfrak{N}' \leq \mathfrak{A}$ hatte sich schon unter 1. ergeben, $\mathfrak{Z}(\mathfrak{N}) \leq \mathfrak{A}$ unter 2. Also ist $\mathfrak{N}'$ zyklisch. Wegen $N^q \in \mathfrak{Z}(\mathfrak{N})$ folgt für beliebige $N_1, N_2 \in \mathfrak{N}$ (ZASSENHAUS [22], S. 57, Formel (9))

$$(N_1, N_2)^q = (N_1^q, N_2) = E$$

und daher $|\mathfrak{N}'| = q$.

4. Ist $q > 2$, so gilt wegen $|\mathfrak{N}'| = q$ die Gleichung (ZASSENHAUS [22], S. 57, Formel (10))

$$(N_1 N_2)^q = (N_2, N_1)^{\binom{q}{2}} N_1^q N_2^q = N_1^q N_2^q.$$

Also bilden die Elemente der Ordnung q von $\mathfrak{N}$ zusammen mit E eine charakteristische Untergruppe $\Omega_1(\mathfrak{N})$ von $\mathfrak{N}$.

Angenommen, $\Omega_1(\mathfrak{N})$ sei abelsch. Dann ist $\mathfrak{N} \neq \Omega_1(\mathfrak{N})$, also nach 2. $\Omega_1(\mathfrak{N}) \leq \mathfrak{Z}(\mathfrak{N})$, und $\Omega_1(\mathfrak{N})$ ist zyklisch. Somit enthält $\mathfrak{N}$ nur eine Untergruppe der Ordnung q. Wegen $q > 2$ ist dann $\mathfrak{N}$ bekanntlich selbst zyklisch (ZASSENHAUS [22], S. 112, Satz 15), im Widerspruch zu $\mathfrak{N}' \neq \mathfrak{E}$.

$\Omega_1(\mathfrak{N})$ ist also invariant in $\mathfrak{G}$ und nichtabelsch, woraus dann $\mathfrak{N} = \Omega_1(\mathfrak{N})$ folgt.

5. Für $q = 2$ bilden wegen $|\mathfrak{N}'| = 2$ und

$$(N_1 N_2)^4 = (N_2, N_1)^{\binom{4}{2}} N_1^4 N_2^4 = N_1^4 N_2^4$$

(ZASSENHAUS [22], S. 57, Formel (10)) die in $\mathfrak{N}$ liegenden Lösungen von $N^4 = E$ eine charakteristische Untergruppe $\Omega_2(\mathfrak{N})$ von $\mathfrak{N}$. Da $\mathfrak{N}$ nicht nur eine Untergruppe der Ordnung 4 enthält (ZASSENHAUS [22], S. 113, Satz 16), ist $\Omega_2(\mathfrak{N})$ nicht abelsch, somit $\Omega_2(\mathfrak{N}) = \mathfrak{N}$.

6. Ist Char. $K = p$, so kann wegen der Irreduzibilität von $\mathfrak{G}$ über K die Ordnung von $\mathfrak{N}$ keine Potenz von p sein (Hilfssatz 1). Also ist $q \neq p$. Falls $\mathfrak{A}$ nur aus Matrizen der Gestalt aE besteht, so auch $\mathfrak{N}'$ nach 3.; für $K = GF(p^f)$ haben wir dann $q = |\mathfrak{N}'|\,|p^f - 1$.

Hilfssatz II (Einbettung von $\mathfrak{N}$ in $\mathfrak{G}$).

Voraussetzung. Wie in Hilfssatz I.

Behauptung.

1. Ist $\mathfrak{N}$ irreduzibel über K und K endlich oder K algebraisch abgeschlossen, so ist der Zentralisator $\mathfrak{C}(\mathfrak{N})$ von $\mathfrak{N}$ in $\mathfrak{G}$ gleich $\mathfrak{A}$.

2. Sei $\mathfrak{C}(\mathfrak{H})$ der Zentralisator des Hauptfaktors $\mathfrak{H}=\mathfrak{N}/\mathfrak{Z}(\mathfrak{N})$ von $\mathfrak{G}$. Ist $|\mathfrak{H}|=q^c$, so ist $\{\mathfrak{C}(\mathfrak{H})\cap\mathfrak{C}(\mathfrak{Z}(\mathfrak{N}))\}/\mathfrak{C}(\mathfrak{N})$ isomorph zu einer Untergruppe des c-fachen direkten Produktes von $\mathfrak{Z}(\mathfrak{N})$ mit sich selbst.

3. Wir fassen die Faktorgruppe $\mathfrak{H}$ als Modul der Dimension c über $GF(q)$ auf. Durch Transformation mit Elementen aus $\mathfrak{C}(\mathfrak{Z}(\mathfrak{N}))$ werden dann auf $\mathfrak{H}$ lineare Abbildungen bewirkt, welche eine nichtausgeartete schiefe Bilinearform mit Werten aus $GF(q)$ invariant lassen. Daher ist $\mathfrak{C}(\mathfrak{Z}(\mathfrak{N}))/\{\mathfrak{C}(\mathfrak{Z}(\mathfrak{N}))\cap\mathfrak{C}(\mathfrak{H})\}$ isomorph zu einer Untergruppe der symplektischen Gruppe $C(2m, q)$ vom Grad $c=2m$ über $GF(q)$.

Beweis.

1. Wegen der Irreduzibilität von $\mathfrak{N}$ über dem Körper K ist $\mathfrak{C}(\mathfrak{N})$ nach dem Schurschen Lemma Untergruppe der multiplikativen Gruppe eines Schiefkörpers S. Ist K endlich, so ist S (als Teilring des vollen Matrizenringes vom Grad n über K) endlich, daher nach dem Satz von WEDDERBURN-WITT (ZASSENHAUS [22], S. 70) kommutativ. Also ist $\mathfrak{C}(\mathfrak{N})$ abelsch. Ist K algebraisch abgeschlossen, so folgt die Behauptung sofort aus der Aussage des Schurschen Lemmas für diesen Fall.

Wegen $\mathfrak{N}\leq\mathfrak{C}(\mathfrak{A})$ gilt $\mathfrak{A}\leq\mathfrak{C}(\mathfrak{N})$. Da aber $\mathfrak{A}$ ein maximaler abelscher Normalteiler ist, folgt nun $\mathfrak{A}=\mathfrak{C}(\mathfrak{N})$.

2. Nach Hilfssatz I.3 ist $\mathfrak{Z}(\mathfrak{N})$ zyklisch; denn $\mathfrak{A}$ ist zyklisch. Sei nun $C\in\mathfrak{C}(\mathfrak{H})\cap\mathfrak{C}(\mathfrak{Z}(\mathfrak{N}))$, Z ein erzeugendes Element von $\mathfrak{Z}(\mathfrak{N})$ und $N_1, \ldots, N_c$ Vertreter einer Basis von $\mathfrak{N}/\mathfrak{Z}(\mathfrak{N})$. Dann gilt

$$C N_i C^{-1} = N_i Z^{a_i} \quad (i = 1, \ldots, c).$$

Durch $C\to(Z^{a_1}, \ldots, Z^{a_c})$ wird dann ein Homomorphismus von $\mathfrak{C}(\mathfrak{H})\cap\mathfrak{C}(\mathfrak{Z}(\mathfrak{N}))$ in das direkte Produkt von c Gruppen vom Typ $\mathfrak{Z}(\mathfrak{N})$ vermittelt. Der Kern dieses Homomorphismus ist offenbar $\mathfrak{C}(\mathfrak{N})$.

3. Ist K ein erzeugendes Element von $\mathfrak{N}'$, so definieren wir eine schiefe Bilinearform $f(N_1, N_2)$ auf $\mathfrak{N}/\mathfrak{Z}(\mathfrak{N})$ durch

$$(N_1, N_2) = K^{f(N_1, N_2)}.$$

Offenbar hängt f nur von den Restklassen der N_i mod $\mathfrak{Z}(\mathfrak{N})$ ab, kann daher als Funktion auf $\mathfrak{N}/\mathfrak{Z}(\mathfrak{N})$ aufgefaßt werden. Wegen $|\mathfrak{N}'|=q$ ist der Wert von f ein Element von $GF(q)$.

Aus $(N_1, N_2)=(N_2, N_1)^{-1}$ folgt $f(N_1, N_2)=-f(N_2, N_1)$, d.h. f ist schief.

Wegen $\mathfrak{N}' \leq \mathfrak{Z}(\mathfrak{N})$ gilt (ZASSENHAUS [22], S. 57, Formel (7a))

$$(N_1 N_2, N_3) = (N_1, N_3)\,(N_2, N_3),$$

also $f(N_1 N_2, N_3) = f(N_1, N_3) + f(N_2, N_3)$, d.h. f ist additiv.

Für $G \in \mathfrak{G}(\mathfrak{Z}(\mathfrak{N}))$ ist wegen $\mathfrak{N}' \leq \mathfrak{Z}(\mathfrak{N})$

$$(N_1^G, N_2^G) = K^{f(N_1^G, N_2^G)} = (N_1, N_2)^G = [K^{f(N_1, N_2)}]^G = K^{f(N_1, N_2)},$$

also $f(N_1^G, N_2^G) = f(N_1, N_2)$, und f ist invariant.

Schließlich ist f nicht ausgeartet, daher bekanntlich c gerade: Aus $f(X, N) = 0$ für alle $N \in \mathfrak{N}$ folgt nämlich $X \in \mathfrak{Z}(\mathfrak{N})$, so daß dem Element X im Modul $\mathfrak{N}/\mathfrak{Z}(\mathfrak{N})$ das Nullelement entspricht.

Da $\mathfrak{N}/\mathfrak{Z}(\mathfrak{N})$ ein Hauptfaktor von $\mathfrak{G}$ ist, muß die von den inneren Automorphismen von $\mathfrak{G}$ auf dem Modul $\mathfrak{H}$ bewirkte Gruppe über $GF(q)$ irreduzibel sein.

Hilfssatz III (Darstellungen von $\mathfrak{N}$).

Voraussetzung. Wie in Hilfssatz I.

Behauptung.

1. Sei $K \subseteq L$, Char. $K \neq q$ und $D(\mathfrak{N})$ eine Darstellung von $\mathfrak{N}$ über L. Dann läßt sich $D(\mathfrak{N})$ in dem Oberkörper J des Primkörpers P von K auf monomiale Gestalt bringen, wenn J für $q > 2$ primitive q-te Einheitswurzeln, für $q = 2$ primitive 4-te Einheitswurzeln enthält.

2. Ist k der Grad einer treuen, absolut irreduziblen Darstellung von $\mathfrak{N}$ über einem Körper L mit Char. $L \neq q$, so gilt $k^2 = |\mathfrak{N}/\mathfrak{Z}(\mathfrak{N})| = q^{2m}$, also $k = q^m$.

3. Ist $D(\mathfrak{N})$ treu und irreduzibel vom Grad k über $K = GF(p^l)$ und $p^l \equiv 1(q)$ für $q > 2$, $p^l \equiv 1(4)$ für $q = 2$, so ist $D(\mathfrak{N})$ absolut irreduzibel und $k = q^m$.

4. Ist $q = 2$ und $|\mathfrak{Z}(\mathfrak{N})| = 2$, so besitzt $\mathfrak{N}$ über jedem Körper L mit Char. $L \neq 2$, welcher primitive 4-te Einheitswurzeln enthält, genau eine absolut irreduzible, treue Darstellung; diese hat den Grad 2^m.

Beweis.

1. Der bekannte Beweis dafür, daß sich jede Darstellung einer q-Gruppe über dem komplexen Zahlkörper C auf monomiale Gestalt bringen läßt (SPEISER [15], S. 196) beruht allein auf der Tatsache, daß sich die Darstellungen jeder abelschen Untergruppe von $\mathfrak{N}$ über C auf Diagonalgestalt bringen lassen.

Ist Char. $K \neq q$, so läßt sich jede Darstellung einer abelschen Untergruppe $\mathfrak{B}$ von $\mathfrak{N}$ auf Diagonalgestalt bringen, und zwar in jedem Körper, welcher alle Eigenwerte der Matrizen von $\mathfrak{B}$ enthält. Wegen 4. bzw. 5. von Hilfssatz I genügt jeder Oberkörper J von P, welcher primitive q-te bzw. primitive 4-te Einheitswurzeln enthält.

2. Wir bestimmen zunächst den Grad einer absolut irreduziblen, treuen Darstellung von $\mathfrak{N}$ über dem komplexen Zahlkörper C. Dabei verwenden wir eine Charakterrechnung von R. FRUCHT (FRUCHT [4], S. 20).

Wegen $\mathfrak{N}' \leq \mathfrak{Z}(\mathfrak{N})$ ist jeder Kommutator (N_1, N_2) von Elementen N_1 und N_2 aus $\mathfrak{N}$ nach dem Schurschen Lemma ein Vielfaches cE der Einheitsmatrix. Für $N \notin \mathfrak{Z}(\mathfrak{N})$ existiert dann ein $Y \in \mathfrak{N}$ mit $YNY^{-1}N^{-1} = cE \neq E$. Dann folgt für die Spur $S(N)$ von N

$$S(N) = S(YNY^{-1}) = S(YNY^{-1}N^{-1}N) = S(cN) = cS(N),$$

wegen $c \neq 1$ also $S(N) = 0$.

Somit haben alle nicht in $\mathfrak{Z}(\mathfrak{N})$ liegenden Elemente von $\mathfrak{N}$ in der betrachteten Darstellung die Spur 0. Aus der bekannten Charakterrelation

$$\sum_{N \in \mathfrak{N}} |S(N)|^2 = |\mathfrak{N}|$$

ergibt sich nun

$$\sum_{Z \in \mathfrak{Z}(\mathfrak{N})} |S(Z)|^2 = |\mathfrak{Z}(\mathfrak{N})|\, k^2 = |\mathfrak{N}|,$$

also $k^2 = |\mathfrak{N}/\mathfrak{Z}(\mathfrak{N})| = q^{2m}$.

Diese Überlegungen gelten nur für den komplexen Zahlkörper. Aber das Ergebnis überträgt sich auf den Fall, daß die Charakteristik von L zur Ordnung von $\mathfrak{N}$ teilerfremd ist (s. BRAUER und NESBITT [1]); denn die irreduziblen Darstellungen von $\mathfrak{N}$ über L entstehen aus geeigneten irreduziblen Darstellungen über algebraischen Zahlkörpern endlichen Grades durch Restklassenbildung nach Primidealteilern $\mathfrak{p}$ von $p = \text{Char. } L$.

3. Wegen $p^f \equiv 1 (q)$ bzw. $p^f \equiv 1 (4)$ enthält $GF(p^f)$ primitive q-te bzw. 4-te Einheitswurzeln. Nach 1. dieses Hilfssatzes lassen sich dann die absolut irreduziblen Darstellungen von $\mathfrak{N}$ bereits in $GF(p^f)$ angeben. Da $D(\mathfrak{N})$ über $GF(p^f)$ irreduzibel ist, ist also $D(\mathfrak{N})$ sogar absolut irreduzibel. Aus 2. dieses Hilfssatzes folgt dann $k = q^m$.

4. Zunächst beschränken wir uns wieder auf den komplexen Zahlkörper C. Aus $|\mathfrak{Z}(\mathfrak{N})| = 2$ folgt $\mathfrak{N}' = \mathfrak{Z}(\mathfrak{N})$. Wegen $|\mathfrak{N}/\mathfrak{Z}(\mathfrak{N})| = 2^{2m}$ hat $\mathfrak{N}$ genau 2^{2m} verschiedene Darstellungen vom Grad 1 über C. Aus der bekannten „Trennungseigenschaft" der absolut irreduziblen Darstellungen folgt, daß es über C eine absolut irreduzible Darstellung $D(\mathfrak{N})$ von $\mathfrak{N}$ gibt, welche $\mathfrak{Z}(\mathfrak{N})$ treu wiedergibt. Dann ist $D(\mathfrak{N})$ treu, da wegen $|\mathfrak{Z}(\mathfrak{N})| = 2$ jeder Normalteiler von $\mathfrak{N}$ die Untergruppe $\mathfrak{Z}(\mathfrak{N})$ enthält. Nach 2. hat $D(\mathfrak{N})$ den Grad 2^m. Die Quadratsumme der Grade der bisher bestimmten Darstellungen beträgt

$$2^{2m} + 2^{2m} = 2^{2m+1} = |\mathfrak{N}|.$$

Damit ist die bekannte Vollständigkeitsrelation erfüllt, unsere Aufstellung der absolut irreduziblen Darstellungen von $\mathfrak{N}$ über C also vollständig.

Wie in 2. überträgt sich dieses Ergebnis auf jeden Körper L mit Char. $L \neq 2$, welcher primitive 4-te Einheitswurzeln enthält.

Nun können wir die Beschreibung des Induktionsverfahrens für den in dieser Arbeit allein interessierenden Fall $\mathfrak{C}(\mathfrak{Z}(\mathfrak{N})) = \mathfrak{G}$ beenden: Ist $2m < n = q^m$,

so können wir auf die zu $\overline{\mathfrak{G}} = \mathfrak{G}/\mathfrak{C}(\mathfrak{H})$ isomorphe Untergruppe des symplektischen Gruppe $C(2m, q)$ unsere Induktionsannahme anwenden. (Es sei ausdrücklich vermerkt, daß wir an dieser Stelle Gruppen über Körpern mit von 0 verschiedener Charakteristik zu betrachten haben, selbst dann, wenn wir nur auf Resultate für den Fall „Charakteristik 0" abzielen.) Nun gilt es im speziellen Fall noch, die Ergebnisse über $\overline{\mathfrak{G}}$ (durch Induktion gewonnen), $\mathfrak{C}(\mathfrak{H})/\mathfrak{C}(\mathfrak{N})$ (aus Hilfssatz II.2) und $\mathfrak{C}(\mathfrak{N})$ (aus Hilfssatz II.1) zu einer Aussage über $\mathfrak{G}$ zusammenzusetzen. Der Sonderfall $2m \gtrsim n = q^m$ tritt nur für $n = 2$ und $n = 4$ ein und muß gesondert behandelt werden.

§ 3. Über die Darstellungen auflösbarer Gruppen

Alle in diesem Paragraphen auftretenden Gruppen $\mathfrak{G}$ seien endlich und auflösbar.

Satz 6. Sei $\mathfrak{G}$ irreduzibel und primitiv vom Grad n über einem algebraisch abgeschlossenen Körper K der Charakteristik 0. Dann gilt $n^2 \mid \mid \mathfrak{G}/\mathfrak{Z}(\mathfrak{G})\mid$.

Beweis durch Induktion nach n.

Da $\mathfrak{G}$ primitiv ist, enthalten die reduziblen Normalteiler von $\mathfrak{G}$ lauter äquivalente irreduzible Bestandteile. Alle abelschen Normalteiler liegen also im Zentrum $\mathfrak{Z}(\mathfrak{G})$ von $\mathfrak{G}$ und bestehen aus Vielfachen von E. Also liegt Fall 3 von § 2 vor. Hat $\mathfrak{G}$ einen minimalen nichtabelschen Normalteiler $\mathfrak{N}$, der über K irreduzibel ist, so gilt nach Hilfssatz III.2 [mit $\mathfrak{Z}(\mathfrak{G})$ statt $\mathfrak{A}$] die Beziehung $n^2 = \mid \mathfrak{N}/\mathfrak{Z}(\mathfrak{N})\mid$, also wegen $\mathfrak{N} \cap \mathfrak{Z}(\mathfrak{G}) = \mathfrak{Z}(\mathfrak{N})$ dann

$$n^2 = \mid \mathfrak{N}/\mathfrak{Z}(\mathfrak{N})\mid = \mid \mathfrak{N}\,\mathfrak{Z}(\mathfrak{G})/\mathfrak{Z}(\mathfrak{G})\mid \mid\mid \mathfrak{G}/\mathfrak{Z}(\mathfrak{G})\mid .$$

Zu untersuchen bleibt noch der Fall, daß alle minimalen nichtabelschen Normalteiler reduzibel sind. Sei $\mathfrak{N}$ ein solcher, der in n_1 Teile $A(\mathfrak{N})$ vom Grad $n_2 > 1$ zerfalle. Wieder gelten für $\mathfrak{N}$ die Ergebnisse von Hilfssatz III und liefern $n_2^2 = \mid \mathfrak{N}/\mathfrak{Z}(\mathfrak{N})\mid$. Der Normalteiler $\mathfrak{M} = \mathfrak{N}\,\mathfrak{Z}(\mathfrak{G})$ von $\mathfrak{G}$ zerfällt in gleicher Weise wie $\mathfrak{N}$, und es ist

$$\mid \mathfrak{M}/\mathfrak{Z}(\mathfrak{G})\mid = \mid \mathfrak{N}/\mathfrak{N} \cap \mathfrak{Z}(\mathfrak{G})\mid = \mid \mathfrak{N}/\mathfrak{Z}(\mathfrak{N})\mid = n_2^2 .$$

Entsprechend dem Zerfallen von $\mathfrak{M}$ haben wir nun nach Satz 3 die Produktzerfegung $\mathfrak{G} = P_1(\mathfrak{G}) \times P_2(\mathfrak{G})$ mit Grad $P_i(\mathfrak{G}) = n_i$. Die zu den $P_i(\mathfrak{G})$ gehörigen Gruppen $\mathfrak{N}_i$ sind absolut irreduzibel und primitiv, da $\mathfrak{G}$ diese Eigenschaften hat; denn einer Einteilung der Darstellungsmoduln $\mathfrak{m}_i$ der $\mathfrak{N}_i$ bzw. $P_i(\mathfrak{G})$ in invariante Teilmoduln oder Imprimitivitätsgebiete würde eine gleichartige Einteilung des Produktmoduls $\mathfrak{m} = \mathfrak{m}_1 \times \mathfrak{m}_2$ entsprechen. Wegen $n_1 < n$ liefert uns nun die Induktionsannahme $n_1^2 \mid \mid \mathfrak{N}_1/\mathfrak{Z}(\mathfrak{N}_1)\mid$. Nun ist aber P_1 eine projektive Darstellung von $\mathfrak{G}/\mathfrak{M}$. Daher gilt $\mid \mathfrak{N}_1/\mathfrak{Z}(\mathfrak{N}_1)\mid \mid \mid \mathfrak{G}/\mathfrak{M}\mid$. Insgesamt erhalten wir

$$n^2 = n_1^2\, n_2^2 \mid \mid \mathfrak{G}/\mathfrak{M}\mid \mid \mathfrak{M}/\mathfrak{Z}(\mathfrak{G})\mid = \mid \mathfrak{G}/\mathfrak{Z}(\mathfrak{G})\mid ,$$

was zu beweisen war.

Es sei noch bemerkt, daß in den Voraussetzungen von Satz 6 weder die Auflösbarkeit noch die Primitivität gestrichen werden können. Daß die Auflösbarkeit unentbehrlich ist, zeigt das Beispiel der treuen Darstellung der Ikosaedergruppe vom Grad 3; daß auf die Primitivität nicht verzichtet werden kann, lehrt die monomiale, absolut irreduzible Darstellung der symmetrischen Gruppe $\mathfrak{S}_3$ vom Grad 2.

Satz 7. Sei $\mathfrak{G}$ irreduzibel über dem algebraisch abgeschlossenen Körper K der Charakteristik p. Dann besitzt $\mathfrak{G}$ ein System von Imprimitivitätsgebieten $\mathfrak{r}_i$, deren Dimension nicht durch p teilbar ist. Ist insbesondere $n = p^a$, so läßt $\mathfrak{G}$ sich auf monomiale Gestalt bringen.

Beweis. Wir führen wieder eine Induktion nach dem Grad n von $\mathfrak{G}$ durch und betrachten nacheinander die drei Fälle von § 2.

Fall 1. Nun ist $\mathfrak{G}$ imprimitiv. Sei $\mathfrak{G}_1$ diejenige Untergruppe, welche $\mathfrak{r}_1$ fest läßt. Dann ist

$$\mathfrak{G}_1 = \begin{pmatrix} B(\mathfrak{G}_1) & 0 & \dots & 0 \\ 0 & & & \\ \vdots & & * & \\ 0 & & & \end{pmatrix}$$

und $B(\mathfrak{G}_1)$ absolut irreduzibel (Satz 2). Ist die Dimension des Moduls $\mathfrak{r}_1$ noch durch p teilbar, so ist nach unserer Induktionsannahme auch $B(\mathfrak{G}_1)$ imprimitiv, und die Imprimitivitätsgebiete von $B(\mathfrak{G}_1)$ geben Anlaß zu einem System von Imprimitivitätsgebieten von $\mathfrak{G}$, welche kleinere Dimension als $\mathfrak{r}_1$ haben. Wählen wir also $\mathfrak{r}_1$ von minimaler Dimension, so ist diese Dimension zu p teilerfremd, was zu zeigen war.

Fall 2. Nun haben wir die Produktzerlegung $\mathfrak{G} = P_1(\mathfrak{G}) \times P_2(\mathfrak{G})$. Nach der Induktionsannahme zerfallen die Darstellungsmoduln $\mathfrak{m}_i$ der $P_i(\mathfrak{G})$ (bzw. der zugehörigen Gruppen $\mathfrak{N}_i$) in Imprimitivitätsgebiete $\mathfrak{r}_j^{(i)}$, deren Dimensionen zu p teilerfremd sind. Dann sind die Produktmoduln $\mathfrak{r}_j^{(1)} \times \mathfrak{r}_k^{(2)}$ Imprimitivitätsgebiete von $\mathfrak{G}$ und haben ebenfalls zu p teilerfremde Dimensionen.

Fall 3. Jetzt ist $\mathfrak{N}$ absolut irreduzibel, $|\mathfrak{N}| = q^b$ und $|\mathfrak{N}/\mathfrak{Z}(\mathfrak{N})| = n^2 = q^{2m}$ (Hilfssatz III. 2). Ist $p \mid n$, so folgt $q = p$. Aber dies ist wegen der Irreduzibilität von $\mathfrak{G}$ nicht möglich, da $\mathfrak{G}$ keinen von $\mathfrak{E}$ verschiedenen Normalteiler von p-Potenzordnung besitzt (Hilfssatz 1).

Damit ist der Beweis von Satz 7 beendet.

Satz 8. Sei $\mathfrak{G}$ irreduzibel vom Grad n über dem algebraisch abgeschlossenen Körper K der Charakteristik O. Für jeden Primteiler p von n gebe es einen abelschen Normalteiler $\mathfrak{A}(p)$ von $\mathfrak{G}$ mit p-überauflösbarer Faktorgruppe[4]) $\mathfrak{G}/\mathfrak{A}(p)$. Dann läßt sich $\mathfrak{G}$ über K auf monomiale Gestalt bringen.

[4]) Eine auflösbare Gruppe heißt p-überauflösbar, wenn alle Hauptfaktoren von p-Potenzordnung zyklisch sind.

Beweis. Wieder verwenden wir unser Induktionsverfahren.

Fall 1. Da sich die Voraussetzung über die Existenz eines abelschen Normalteilers mit p-überauflösbarer Faktorgruppe auf alle Untergruppen von $\mathfrak{G}$ überträgt, kann die über K irreduzible Darstellung $B(\mathfrak{G}_1)$ nach der Induktionsannahme monomial gemacht werden. Dann kann aber nach der üblichen Schlußweise (SPEISER [15], S. 197) auch ganz $\mathfrak{G}$ monomial gemacht werden.

Weiterhin können wir also annehmen, daß alle abelschen Normalteiler von $\mathfrak{G}$ in lauter äquivalente Teile vom Grad 1 zerfallen, daher im Zentrum $\mathfrak{Z}(\mathfrak{G})$ von $\mathfrak{G}$ liegen. Daraus folgt, daß für die betrachteten p sogar $\mathfrak{G}$ selbst p-überauflösbar ist.

Fall 2. Nun ist $\mathfrak{G} = P_1(\mathfrak{G}) \times P_2(\mathfrak{G})$. Die $P_i(\mathfrak{G})$ entstehen aus geeigneten Gruppen $\mathfrak{N}_i$. Da aber die $\mathfrak{N}_i$ durch zentrale Erweiterung von zyklischen Gruppen mit homomorphen Bildern von $\mathfrak{G}$ gebildet sind, erfüllen auch die $\mathfrak{N}_i$ unsere Voraussetzung der p-Überauflösbarkeit bezüglich der p mit $p|n$. Wegen der absoluten Irreduzibilität von $\mathfrak{G}$ sind auch die $P_i(\mathfrak{G})$ absolut irreduzibel, dann aber auch die $\mathfrak{N}_i$. Da die Grade n_i der $P_i(\mathfrak{G})$ bzw. der $\mathfrak{N}_i$ Teiler von n sind, können wir unsere Induktionsannahme auf die $\mathfrak{N}_i$ anwenden und sehen, daß diese monomial gemacht werden können. Das gleiche gilt dann auch für die $P_i(\mathfrak{G})$ und deren Kronecker-Produkt $\mathfrak{G}$.

Fall 3. Nach Hilfssatz III.2 ist nun $|\mathfrak{N}/\mathfrak{Z}(\mathfrak{N})| = n^2 = q^{2m}$. Daher tritt q als einziger Primteiler von n auf. Wegen unserer verschärften Voraussetzung ist also $\mathfrak{G}$ q-überauflösbar. Da $\mathfrak{N}/\mathfrak{Z}(\mathfrak{N})$ ein minimaler Normalteiler von $\mathfrak{G}/\mathfrak{Z}(\mathfrak{N})$ der Ordnung q^{2m} ist (Hilfssatz I.2), folgt $2m = 1$. Dies ist aber unmöglich. Also kann unter den vorliegenden Umständen Fall 3 nicht eintreten.

Damit haben wir bewiesen, daß sich $\mathfrak{G}$ über K auf monomiale Gestalt bringen läßt.

Satz 8 ist eine Verallgemeinerung des folgenden Ergebnisses von ZASSENHAUS (ZASSENHAUS [20], S. 207, Satz 10): Hat $\mathfrak{G}$ einen abelschen Normalteiler mit überauflösbarer Faktorgruppe, so läßt sich jede Darstellung von $\mathfrak{G}$ über dem komplexen Zahlkörper auf monomiale Gestalt bringen.

§ 4. Abschätzung der Stufe

ZASSENHAUS hat folgenden Satz bewiesen (ZASSENHAUS [21], Satz 7): Sei $\mathfrak{G}$ eine auflösbare lineare Gruppe vom Grad n. Dann liegt die Stufe von $\mathfrak{G}$ unterhalb einer nur von n abhängigen Schranke $k(n)$.

Der Beweis von ZASSENHAUS läßt sich zu einer Bestimmung von $k(n)$ ausbauen; jedoch wird die so erhaltene Schranke sehr schlecht, da der letzte Beweisschritt bei ZASSENHAUS die Abschätzung der Stufe aller auflösbaren, primitiven, unimodularen Gruppen $\mathfrak{G}$ vom Grad n mit

$$|\mathfrak{G}| \leq (n^{n^2} + 1)!\, n$$

erfordert. Wir geben ein wesentlich schärferes Ergebnis an:

Satz 9. Sei $\mathfrak{G}$ auflösbar vom Grad n. Dann ist $\mathfrak{G}^{(2n)} = \mathfrak{E}$.

Beweis. Wir verwenden wieder das Induktionsprinzip von § 2. Dabei unterscheiden sich die ersten Schritte nicht wesentlich von dem Beweisgang bei ZASSENHAUS; erst in d) weichen wir entscheidend ab.

In unsere Aussage geht der Körper K, dem die Elemente der Matrizen von $\mathfrak{G}$ entnommen sind, nicht ein, so daß wir zur algebraisch abgeschlossenen Hülle von K übergehen dürfen.

a) Zunächst sei $\mathfrak{G}$ reduzibel, also in geeigneten Koordinaten

$$\mathfrak{G} = \begin{pmatrix} A_{11}(\mathfrak{G}) & 0 \\ A_{21}(\mathfrak{G}) & A_{22}(\mathfrak{G}) \end{pmatrix}.$$

Ist Grad $A_{ii}(\mathfrak{G}) = n_i$ mit $n_1 + n_2 = n$, so liefert unsere Induktionsannahme

$$A_{ii}(\mathfrak{G})^{(2n_i)} = E_{n_i}.$$

Falls $n_1 \geq n_2$, so besteht $\mathfrak{G}^{(2n_1)}$ aus Matrizen der Gestalt

$$\begin{pmatrix} E_{n_1} & 0 \\ A & E_{n_2} \end{pmatrix},$$

und ist daher abelsch. Es folgt $\mathfrak{G}^{(2n_1+1)} = \mathfrak{E}$ und dann unter Beachtung von $n_1 < n$ auch $\mathfrak{G}^{(2n)} = \mathfrak{E}$.

Weiterhin können wir also annehmen, daß $\mathfrak{G}$ absolut irreduzibel ist, so daß nun die Fallunterscheidung von § 2 in Kraft tritt.

b) Es liege Fall 1 von § 2 vor, $\mathfrak{G}$ sei imprimitiv mit k Imprimitivitätsgebieten der Dimension m. Unsere Induktionsannahme zeigt jetzt $\mathfrak{D}^{(2m)} = \mathfrak{E}$. Ist $k < n$, so können wir auch auf $\mathfrak{P}_k \cong \mathfrak{G}/\mathfrak{D}$ die Induktionsannahme anwenden und erhalten $\mathfrak{P}_k^{(2k)} = \mathfrak{E}$. Insgesamt ist also $\mathfrak{G}^{(2m+2k)} = \mathfrak{E}$. Wegen $km = n$ und $1 < k < n$ folgt nun $k + m \leq n/2 + n/2 = n$, womit wir in diesem Falle fertig sind.

Es bleibt noch der Fall $k = n$ zu behandeln. Dann besteht $\mathfrak{D}$ aus Diagonalmatrizen, ist daher abelsch. Bei geeigneter Wahl der Koordinaten können wir $\mathfrak{P}_n$ über dem komplexen Zahlkörper C vollständig reduzieren in

$$\begin{pmatrix} 1 & 0 \\ 0 & B_{22}(\mathfrak{P}_n) \end{pmatrix}.$$

Die Anwendung der Induktionsannahme auf die Gruppe $B_{22}(\mathfrak{P}_n)$ vom Grade $n-1$ liefert uns nun $B_{22}(\mathfrak{P}_n)^{(2n-2)} = \mathfrak{E}$, also auch $\mathfrak{P}_n^{(2n-2)} = \mathfrak{E}$. Wegen $\mathfrak{D}' = \mathfrak{E}$ folgt dann das gewünschte Resultat $\mathfrak{G}^{(2n)} = \mathfrak{E}$.

c) Jetzt betrachten wir Fall 2 mit der Produktzerlegung $\mathfrak{G} = P_1(\mathfrak{G}) \times P_2(\mathfrak{G})$ und Grad $P_i(\mathfrak{G}) = n_i > 1$. Auf die zu den $P_i(\mathfrak{G})$ bei geeignet normierten Faktorensystemen gehörigen Gruppen $\mathfrak{A}_i$ können wir wieder unsere Induktionsannahme anwenden. Ist etwa $n_1 \geq n_2$, so erhalten wir

$$\mathfrak{A}_i^{(2n_i)} = \{E_{n_i}\}.$$

und dann, da mit dem Kronecker-Produkt komponentenweise gerechnet wird (unter Beachtung einer eventuellen Rücknormierung des Faktorensystems) $\mathfrak{G}^{(2n_1)} = \{aE_n\}$, also $\mathfrak{G}^{(2n_1+1)} = \mathfrak{E}$ und $\mathfrak{G}^{(2n)} = \mathfrak{E}$.

d) Zu behandeln bleibt noch der Fall 3. Jetzt ist $\mathfrak{G}$ primitiv. Nach ZASSENHAUS [21] (Satz 6) bilden dann die Elemente von $\mathfrak{G}$ mit Determinante 1 einen endlichen Normalteiler $\mathfrak{S}$ von $\mathfrak{G}$ mit abelscher Faktorgruppe. Ist $\mathfrak{S}$ abelsch, so ist $\mathfrak{G}'' = \mathfrak{E}$ und unsere Behauptung offenbar richtig. Ist aber $\mathfrak{S}' \neq \mathfrak{E}$, so können wir nach b) und c) annehmen, daß auch $\mathfrak{S}$ absolut irreduzibel ist. Nun betrachten wir einen in $\mathfrak{S}$ liegenden, minimalen nichtabelschen Normalteiler $\mathfrak{N}$ von $\mathfrak{G}$. $\mathfrak{N}$ ist endlich. Nach b) und c) dürfen wir annehmen, daß auch $\mathfrak{N}$ absolut irreduzibel ist. Daher können wir die Hilfssätze II und III anwenden, und zwar mit $\mathfrak{G} = \mathfrak{C}(\mathfrak{Z}(\mathfrak{N}))$.

Ist $2m < n = q^m$, so können wir auf die Gruppe $\overline{\mathfrak{G}} = \mathfrak{G}/\mathfrak{C}(\mathfrak{H})$ vom Grade $2m$ die Induktionsannahme anwenden. Wir bekommen so $\overline{\mathfrak{G}}^{(4m)} = \overline{\mathfrak{E}}$ und dann, da $\mathfrak{C}(\mathfrak{H})/\mathfrak{C}(\mathfrak{N})$ und $\mathfrak{C}(\mathfrak{N})$ abelsch sind (Hilfssatz II. 2 und II. 1) $\mathfrak{G}^{(4m+2)} = \mathfrak{E}$. Für $n \neq 2, 4$ ist aber $4m + 2 \leq 2n = 2q^m$, womit unsere Behauptung in diesem Falle bewiesen ist.

Es bleiben noch die (auch schon durch die Bedingung $2m < q^m$ ausgeschlossenen) Fälle $n = 2$ und $n = 4$ zu betrachten.

Für $n = 2$ ist $\overline{\mathfrak{G}} \leq GL(2, 2)$, daher $\overline{\mathfrak{G}}'' = \overline{\mathfrak{E}}$ und dann $\mathfrak{G}^{(4)} = \mathfrak{E}$. Etwas schwieriger ist der Fall $n = 4$. Dann ist $\overline{\mathfrak{G}}$ eine über $GF(2)$ irreduzible Untergruppe von $GL(4, 2)$ [sogar von $C(4, 2)$]. Das einzige Vielfache der Einheitsmatrix E, welches in $\overline{\mathfrak{G}}$ vorkommen kann, ist E selbst. Somit kann Fall 3 für $\overline{\mathfrak{G}}$ sicher nicht eintreten.

Ist $\overline{\mathfrak{G}}$ reduzibel über einem Erweiterungskörper L von $GF(2)$, so zerfällt es bei geeigneter Wahl von L nach Hilfssatz 3 vollständig in Bestandteile gleichen Grades, die also den Grad 1 oder 2 haben. Die Induktionsannahme liefert uns nun $\overline{\mathfrak{G}}^{(4)} = \overline{\mathfrak{E}}$ und dann $\mathfrak{G}^{(6)} = \mathfrak{E}$.

Nun trete für $\overline{\mathfrak{G}}$ der Fall 1 ein, d. h. $\overline{\mathfrak{G}}$ sei imprimitiv über einem geeigneten Erweiterungskörper von $GF(2)$. Ist $\overline{\mathfrak{G}}$ monomial, so ist $\overline{\mathfrak{P}}_4$ eine Untergruppe der symmetrischen Gruppe $\mathfrak{S}_4$, und wir erhalten $\overline{\mathfrak{P}}_4^{(3)} = \overline{\mathfrak{E}}$, dann $\overline{\mathfrak{G}}^{(4)} = \overline{\mathfrak{E}}$ und daher $\mathfrak{G}^{(6)} = \mathfrak{E}$. Haben die Imprimitivitätsgebiete von $\overline{\mathfrak{G}}$ den Grad 2, so ist $k = 2$, also $\overline{\mathfrak{P}}_2' = \overline{\mathfrak{E}}$ und $\overline{\mathfrak{G}}' \leq \overline{\mathfrak{D}}$. Da $\overline{\mathfrak{D}}$ vollständig in zwei Teile vom Grad 2 zerfällt, diese aber höchstens die Stufe 4 haben, erhalten wir $\overline{\mathfrak{G}}^{(5)} \leq \overline{\mathfrak{D}}^{(4)} = \overline{\mathfrak{E}}$ und dann $\mathfrak{G}^{(7)} = \mathfrak{E}$.

Läßt schließlich $\overline{\mathfrak{G}}$ sich zerlegen in ein Kronecker-Produkt $\overline{\mathfrak{G}} = P_1(\overline{\mathfrak{G}}) \times P_2(\overline{\mathfrak{G}})$ (Fall 2), so ist $\overline{\mathfrak{G}}^{(5)} = \overline{\mathfrak{E}}$ und $\mathfrak{G}^{(7)} = \mathfrak{E}$.

Da Fall 3 für $\overline{\mathfrak{G}}$ nicht eintreten kann, haben wir damit auch für $n = 4$ unsere Behauptung bewiesen.

Für $n = 1$ ist Satz 9 offenbar nicht scharf; aber für $n = 2$ gibt er die genaue Schranke an, wie folgendes Beispiel zeigt:

Die Matrizen

$$A = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}, \quad B = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \quad C = \begin{pmatrix} 1 & -i \\ i & -1 \end{pmatrix}, \quad D = \begin{pmatrix} 1 & -i \\ 1 & i \end{pmatrix}$$

mit den Relationen

$$BAB^{-1}=A^{-1},\ CAC^{-1}=B,\ CBC^{-1}=A,\ DAD^{-1}=AB,$$
$$DBD^{-1}=A,\ A^2=B^2=-E,\ C^2=2E,\ D^3=2(1-i)E,$$
$$CDC^{-1}=\frac{1-i}{2}BD^2$$

liefern über jedem Körper K, welcher eine primitive 4-te Einheitswurzel i enthält, eine Gruppe vom Grad 2 und der Stufe 4. Allerdings ist diese Gruppe nur dann von endlicher Ordnung, wenn K nicht die Charakteristik 0 hat.

Für große n läßt sich Satz 9 wesentlich verschärfen.

Satz 10. Sei $\mathfrak{G}$ auflösbar vom Grad $n>1$. Ist $\mathfrak{G}$ vollständig reduzibel, so gilt $k(\mathfrak{G})\leq 6\,{}_2\log n$, für beliebiges auflösbares $\mathfrak{G}$ $k(\mathfrak{G})\leq 7\,{}_2\log n+1$.

Beweis. Der Beweis weicht nur stellenweise vom Beweis zu Satz 9 ab. Wir fassen uns daher kurz und geben nur die Änderungen an.

a) Ist $\mathfrak{G}$ vollständig reduzibel in absolut irreduzible Teile der Grade $n_i<n$, so können wir auf die Bestandteile mit $n_i>1$ die Induktionsannahme anwenden und erhalten

$$k(\mathfrak{G})\leq \mathrm{Max}\{6\,{}_2\log n_i, 1\}\leq 6\,{}_2\log n.$$

Ist $\mathfrak{G}$ nicht vollständig reduzibel, so wenden wir auf die irreduziblen Teile das bisherige Ergebnis an. Dann sehen wir, daß für $k=[6\,{}_2\log n]$ die Gruppe $\mathfrak{G}^{(k)}$ (in geeigneten Koordinaten) aus Matrizen der Gestalt

$$\begin{pmatrix} E_{n_1} & & \dots & 0 \\ & E_{n_2} & \dots & 0 \\ & & & \vdots \\ * & & & E_{n_l} \end{pmatrix}$$

besteht, also eine Untergruppe der Gruppe aller Dreiecksmatrizen ist. Daher hat $\mathfrak{G}^{(k)}$ höchstens die Stufe $[{}_2\log n]+1$ (GRÜN [6], Satz 8). Wir erhalten somit

$$k(\mathfrak{G})\leq 7\,{}_2\log n+1.$$

Im folgenden können wir uns auf den Fall beschränken, daß $\mathfrak{G}$ absolut irreduzibel ist.

b) Sei wieder k die Anzahl der Imprimitivitätsgebiete von $\mathfrak{G}$. Ist $k\neq n$, so können wir die Induktionsannahme auf die irreduziblen Bestandteile von $\mathfrak{D}$ anwenden; denn als Normalteiler der irreduziblen Gruppe $\mathfrak{G}$ ist $\mathfrak{D}$ nach Satz 1 vollständig reduzibel. Die Permutationsgruppe $\mathfrak{P}_k\cong\mathfrak{G}/\mathfrak{D}$ fassen wir als lineare Gruppe vom Grad k über dem komplexen Zahlkörper auf; als solche ist sie dann vollständig reduzibel. Also dürfen wir die Induktionsannahme auch auf $\mathfrak{P}_k$ anwenden und erhalten $k(\mathfrak{G})\leq 6({}_2\log k+{}_2\log n/k)=6\,{}_2\log n$, wie behauptet.

Für $k = n$ müssen wir den früheren Schluß abändern. Wir dürfen annehmen, daß die Permutationsgruppe $\mathfrak{P}_k$ primitiv ist; denn andernfalls könnten wir Imprimitivitätsgebiete von $\mathfrak{G}$ der Dimension > 1 auffinden, womit dann $k < n$ wäre und der eben behandelte Fall vorläge.

Als primitive, auflösbare Gruppe hat $\mathfrak{P}_k$ bekanntlich Primzahlpotenzgrad p^m und ist Produkt eines elementar abelschen Normalteilers $\mathfrak{T}$ mit einer Untergruppe $\mathfrak{U}$ der linearen Gruppe $GL(m, p)$. Dabei ist $\mathfrak{U}$ irreduzibel über $GF(p)$, läßt sich also über einem genügend großen separablen Erweiterungskörper von $GF(p)$ vollständig in absolut irreduzible Teile zerlegen (Hilfssatz 3).

Ist $m = 1$, so erhalten wir $\mathfrak{U}' = \mathfrak{E}$, daher $\mathfrak{P}_n'' = \mathfrak{E}$ und $\mathfrak{G}''' = \mathfrak{E}$. Dann ist unsere Behauptung wegen $3 < 6 \, {}_2\log 2 \leq 6 \, {}_2\log n$ bewiesen.

Für $m > 1$ können wir auf $\mathfrak{U}$ die Induktionsannahme anwenden und bekommen $k(\mathfrak{U}) \leq 6 \, {}_2\log m$ und daraus $k(\mathfrak{G}) \leq 6 \, {}_2\log m + 2$. Hinreichend für die Richtigkeit unserer Behauptung ist in diesem Falle also die Gültigkeit der Ungleichung

$$(*) \qquad 6 \, {}_2\log m + 2 \leq 6 \, {}_2\log 2^m = 6m.$$

Unter d) werden wir für $m \geq 3$ eine noch schärfere Ungleichung beweisen, aus der (*) folgt.

Für $m = 2$ ist nach Satz 9 $\mathfrak{U}^{(4)} = \mathfrak{E}$, daher $\mathfrak{G}^{(6)} = \mathfrak{E}$. Aus $6 < 6 \, {}_2\log 2^2 \leq 6 \, {}_2\log n$ folgt nun die Richtigkeit unserer Behauptung für $m = 2$.

c) Aus $\mathfrak{G} = P_1(\mathfrak{G}) \times P_2(\mathfrak{G})$ mit Grad $P_i(\mathfrak{G}) = n_i < n$ folgt für $n_1 \geq n_2$ mit $k = 6 \, {}_2\log n_1$ nach der Induktionsannahme $\mathfrak{R}_i^{(k)} = \mathfrak{E}$ und dann $\mathfrak{G}^{(k)} = \{aE_n\}$, also $\mathfrak{G}^{(k+1)} = \mathfrak{E}$. Mit

$$k + 1 \leq 6 \, {}_2\log n_1 + 1 \leq 6 \, {}_2\log n_1 + 6 \, {}_2\log n_2 = 6 \, {}_2\log n$$

ergibt sich nun die Behauptung.

d) Der Fall $\mathfrak{S}' = \mathfrak{E}$ ist wie in Satz 9 trivial. Für $n \neq 2, 4$ können wir wegen $n > 2m \geq 2$ die Induktionsannahme auf die über $GF(q)$ irreduzible Gruppe $\overline{\mathfrak{G}}$ anwenden. Dies liefert $k(\overline{\mathfrak{G}}) \leq 6 \, {}_2\log 2m$ und damit dann $k(\mathfrak{G}) \leq 6 \, {}_2\log 2m + 2$. Dabei gilt $n = q^m \geq 2^m$. Unsere Behauptung ist nun gesichert, wenn wir die Ungleichung

$$(**) \qquad 6 \, {}_2\log 2m + 2 = 6 \, {}_2\log m + 8 \leq 6 \, {}_2\log 2^m = 6m$$

nachweisen. Mit (**) ist sicher auch (*) erfüllt.

(**) ist für $m \geq 3$ erfüllt; denn die Funktion $f(x) = x - {}_2\log x$ steigt für $x \geq 3$ monoton und hat für $x = 3$ einen Wert $> 4/3$.

Nun betrachten wir noch gesondert die Fälle $m = 1$ und $m = 2$, für welche (**) nicht erfüllt ist. Für $q = 2$ erfassen wir dabei die schon zu Beginn von d) ausgeschlossenen Grade $n = 2$ und $n = 4$ mit.

Für $m = 1$ gilt nach Satz 9 $k(\overline{\mathfrak{G}}) \leq 4$, daher $k(\mathfrak{G}) \leq 6$. Die Behauptung ist dann wegen $6 = 6 \, {}_2\log 2 \leq 6 \, {}_2\log n$ richtig.

Für $m=2$ haben wir nach Satz 9 die Abschätzung $k(\overline{\mathfrak{G}}) \leq 8$ und dann $k(\mathfrak{G}) \leq 10$. Aus $10 < 6\,{}_2\log 2^2 \leq 6\,{}_2\log n$ ergibt sich nun die Behauptung.

Damit ist Satz 10 vollständig bewiesen.

Die in Satz 10 angegebene Abschätzung von $k(\mathfrak{G})$ ist der Größenordnung nach die bestmögliche. Denn die Gruppe der Dreiecksmatrizen hat nach GRÜN [6] (Satz 8) die Stufe $[{}_2\log n]+1$ für $n \neq 2^m$ bzw. ${}_2\log n$ für $n=2^m$.

Unter spezielleren Annahmen über $\mathfrak{G}$ können wir Satz 9 noch verschärfen.

Satz 11. $\mathfrak{G}$ sei eine auflösbare Gruppe vom Grad n über dem Körper K. Ist $D(\mathfrak{G})$ eine beliebige Darstellung von $\mathfrak{G}$ über dem komplexen Zahlkörper C oder der algebraisch abgeschlossenen Hülle von K, so lasse $D(\mathfrak{G})$ sich auf monomiale Gestalt bringen. Dann gilt $\mathfrak{G}^{(n)} = \mathfrak{E}$.

Beweis. Wir führen den Beweis durch Induktion nach dem Grad n. Nach unserer Voraussetzung ist $\mathfrak{G}$ monomial. Sei $\mathfrak{D}$ der aus den Diagonalmatrizen von $\mathfrak{G}$ bestehende Normalteiler von $\mathfrak{G}$. Dann ist $\mathfrak{G}/\mathfrak{D}$ isomorph zu einer Permutationsgruppe $\mathfrak{P}_n$ vom Grad n, die wir als lineare Gruppe vom Grad n über C auffassen wollen. $\mathfrak{P}_n$ enthält mindestens einmal die identische Darstellung und ist vollständig reduzibel, läßt sich daher auf die Gestalt

$$\begin{pmatrix} 1 & 0 \\ 0 & \mathfrak{L}_{n-1} \end{pmatrix}$$

bringen. Die Gruppe $\mathfrak{L}_{n-1}$ ist isomorph zu $\mathfrak{P}_n$, also homomorphes Bild von $\mathfrak{G}$ und erfüllt daher unsere Voraussetzungen. Nach der Induktionsannahme ist dann $\mathfrak{P}_n^{(n-1)} = \mathfrak{E}$, also $\mathfrak{G}^{(n-1)} \leq \mathfrak{D}$ und dann $\mathfrak{G}^{(n)} = \mathfrak{E}$, wie behauptet.

Die Voraussetzung in Satz 11, daß auch alle Darstellungen von $\mathfrak{G}$ über C, und nicht nur über K, auf monomiale Gestalt gebracht werden können, läßt sich umgehen, etwa unter Verwendung des folgenden Hilfssatzes von H. WIELANDT (mündliche Mitteilung): Besitzt die Permutationsgruppe $\mathfrak{P}$ über K den identischen Bestandteil genau einmal und nur einen absolut irreduziblen weiteren Bestandteil vom Grad >1, so ist sie vollständig reduzibel. (Ist $K=C$, so ist $\mathfrak{P}$ bekanntlich stets vollständig reduzibel, unabhängig von den Graden der Bestandteile.)

Die Voraussetzungen von Satz 11 treffen z.B. dann zu, wenn $\mathfrak{G}$ lauter abelsche Sylow-Gruppen hat und die Ordnung von $\mathfrak{G}$ zu Char. K teilerfremd ist (N. ITÔ [11]).

§ 5. Abschätzung des Ranges

Eine Abschätzung des Ranges $r(\mathfrak{G})$ einer beliebigen linearen, auflösbaren Gruppe $\mathfrak{G}$ durch den Grad n allein ist bei beliebigem Körper K nicht möglich; denn die Gruppe $\mathfrak{G}$ der Dreiecksmatrizen

$$\begin{pmatrix} a & 0 \\ c & b \end{pmatrix}$$

mit $a, b, c \in GF(p^l)$ und $ab \neq 0$ besitzt die Gruppe der Matrizen

$$\begin{pmatrix} 1 & 0 \\ c & 1 \end{pmatrix}$$

als minimalen Normalteiler. Da dieser Normalteiler f erzeugende Elemente hat, ist $r(\mathfrak{G}) = f$. Dieses f kann aber völlig unabhängig vom Grad $n = 2$ gewählt werden, so daß eine Ungleichung der Gestalt $r(\mathfrak{G}) \leq h(n)$ für keine Funktion $h(n)$ bestehen kann.

Wir wollen nun zeigen, daß wenigstens bei vollständig reduziblen Gruppen gegebenen Grades eine Schranke für den Rang existiert. Zwei Hilfssätze schicken wir voraus:

Hilfssatz 4. Sei $\mathfrak{G}$ eine imprimitive Gruppe des Grades n mit k Imprimitivitätsgebieten $\mathfrak{r}_i$ der Dimension m. Sei $\mathfrak{D}$ derjenige Normalteiler von $\mathfrak{G}$, welcher jedes Imprimitivitätsgebiet $\mathfrak{r}_i$ fest läßt. Ist dann

$$D = \begin{pmatrix} D_1 & 0 & \dots & 0 \\ 0 & D_2 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & D_k \end{pmatrix}$$

eine Matrix aus $\mathfrak{D}$, so nehmen wir zu $\mathfrak{G}$ alle die Matrizen

$$D_{(i)} = \begin{pmatrix} E & 0 & & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots D_i \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & E \end{pmatrix}$$

hinzu, und zwar machen wir dies für jedes $D \in \mathfrak{D}$. Es entsteht dann eine Gruppe $\mathfrak{G}^* \geq \mathfrak{G}$, welche die von den $D_{(i)}$ gebildete Untergruppe $\mathfrak{D}^*$ als Normalteiler besitzt. Dabei ist $\mathfrak{G}^*/\mathfrak{D}^* \cong \mathfrak{G}/\mathfrak{D}$, und diese beiden Faktorgruppen haben gemeinsame Vertreter.

Beweis. Ist

$$GDG^{-1} = D' = \begin{pmatrix} D_1' & 0 & \dots & 0 \\ 0 & D_2' & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & D_k' \end{pmatrix}$$

und $G\mathfrak{r}_i = \mathfrak{r}_{i'}$, so ist $GD_{(i)}G^{-1} = D'_{(i')} \in \mathfrak{D}^*$. Daher bilden die $D_{(i)}$ einen Normalteiler $\mathfrak{D}^*$ der erweiterten Gruppe $\mathfrak{G}^*$. Der Rest der Behauptung ist trivial.

Hilfssatz 5. Wir betrachten die in Hilfssatz 4 konstruierte imprimitive Obergruppe $\mathfrak{G}^*$ von $\mathfrak{G}$. Sei $\mathfrak{G}_1^*$ diejenige Untergruppe von $\mathfrak{G}^*$, welche das erste Imprimitivitätsgebiet $\mathfrak{r}_1$ in sich überführt; $\mathfrak{G}_1^*$ bewirke auf $\mathfrak{r}_1$ die Darstellung $B(\mathfrak{G}_1^*)$. Ist dann $r\left(B(\mathfrak{G}_1^*)\right) \leq c$, so folgt $r_{\mathfrak{G}^*}(\mathfrak{D}^*) \leq ck$.

Beweis. $\mathfrak{D}^*$ enthält nach unserer Konstruktion einen Normalteiler $\mathfrak{D}_1^*$ der Gestalt

$$\mathfrak{D}_1^* = \begin{pmatrix} E & \dots & 0 \\ \vdots & & \vdots \\ 0 & D_1(\mathfrak{D}_1^*) & 0 \\ \vdots & & \vdots \\ 0 & \dots & E \end{pmatrix}.$$

Offenbar ist $\mathfrak{D}^* = \mathfrak{D}_1^* \times \cdots \times \mathfrak{D}_k^*$. Da sich beim Übergang von $\mathfrak{G}$ zu $\mathfrak{G}^*$ die ersten Komponenten D_1 nicht geändert haben, ist wegen $\mathfrak{D} \trianglelefteq \mathfrak{G}_1$ auch $\mathfrak{D}_1^* \trianglelefteq \mathfrak{G}_1^*$.

Sei nun $\mathfrak{E} = \mathfrak{K}_0 \lhd \mathfrak{K}_1 \lhd \cdots \lhd \mathfrak{K}_s = D_1(\mathfrak{D}_1^*)$ das Stück einer Hauptreihe von $B(\mathfrak{G}_1^*)$ unterhalb $D_1(\mathfrak{D}_1^*)$. Aus unserer Voraussetzung $r\left(B(\mathfrak{G}_1^*)\right) \leq c$ folgt $|\mathfrak{K}_{j+1}/\mathfrak{K}_j| = p_j^{a_j}$ mit $a_j \leq c$. Da $\mathfrak{D}_1^*$ durch $D_1(\mathfrak{D}_1^*)$ treu dargestellt wird, entsprechen den $\mathfrak{K}_j$ isomorphe Normalteiler $\mathfrak{M}_j$ von $\mathfrak{G}_1^*$ mit $|\mathfrak{M}_{j+1}/\mathfrak{M}_j| = p_j^{a_j}$.

Die Elemente $G_1, G_2, \dots, G_k$ seien Vertreter von $\mathfrak{G}_1^*$ in $\mathfrak{G}^*$ mit $G_i \mathfrak{r}_i = \mathfrak{r}_1$. Da $\mathfrak{D}_1^*$ nur auf $\mathfrak{r}_1$ von der Identität verschiedene Abbildungen bewirkt, liefert $\mathfrak{D}_1^{*G_i}$ auf allen $\mathfrak{r}_j$ mit $j \neq i$ die identische Abbildung. Somit ist

$$\mathfrak{D}_1^{*G_i} \cap \mathfrak{D}_1^{*G_j} = \mathfrak{E}$$

für $i \neq j$. Nun bilden wir die direkten Produkte

$$\mathfrak{L}_j = \mathfrak{M}_j^{G_1} \times \mathfrak{M}_j^{G_2} \times \cdots \times \mathfrak{M}_j^{G_k}.$$

Die $\mathfrak{L}_j$ sind Normalteiler von $\mathfrak{G}^*$: Sei nämlich G ein beliebiges Element von $\mathfrak{G}^*$. Zu jedem s gibt es $\overline{G_1} \in \mathfrak{G}_1^*$ und t mit $G_s G = \overline{G_1} G_t$. Nun erhalten wir unter Beachtung von

$$\mathfrak{M}_j^{\overline{G_1}} = \mathfrak{M}_j$$

die Aussage

$$\mathfrak{M}_j^{G_s G} = \mathfrak{M}_j^{\overline{G_1} G_t} = \mathfrak{M}_j^{G_t}.$$

Bei Transformation mit G werden also die Faktoren von $\mathfrak{L}_j$ vertauscht und $\mathfrak{L}_j$ selbst bleibt invariant. Aus $|\mathfrak{L}_{j+1}/\mathfrak{L}_j| = |\mathfrak{M}_{j+1}/\mathfrak{M}_j|^k = p_j^{a_j k}$ mit $a_j k \leq ck$ ergibt sich nun wegen $\mathfrak{L}_s = \mathfrak{D}_1^* \times \cdots \times \mathfrak{D}_k^* = \mathfrak{D}^*$ die Behauptung $r_{\mathfrak{G}^*}(\mathfrak{D}^*) \leq ck$.

Wir kommen nun zu der angekündigten Abschätzung von $r(\mathfrak{G})$.

Satz 12. Sei $\mathfrak{G}$ eine endliche, auflösbare Gruppe vom Grad n, welche über dem algebraisch abgeschlossenen Körper K vollständig reduzibel ist. Dann ist $r(\mathfrak{G}) \leq n$.

Beweis. Wieder führen wir nach dem Vorbild von § 2 einen Induktionsbeweis nach n.

a) Zunächst sei $\mathfrak{G}$ reduzibel. Nach unserer Voraussetzung ist dann $\mathfrak{G}$ vollständig reduzibel, also bei passender Wahl des Koordinatensystems

$$\mathfrak{G} = \begin{pmatrix} A_1(\mathfrak{G}) & 0 & \dots & 0 \\ 0 & A_2(\mathfrak{G}) & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & A_j(\mathfrak{G}) \end{pmatrix}$$

mit absolut irreduziblen Darstellungen $A_i(\mathfrak{G})$ der Grade n_i. Auf die $A_i(\mathfrak{G})$ können wir nun unsere Induktionsannahme anwenden und erhalten $r(A_i(\mathfrak{G})) \leq n_i < n$. Da aber $\mathfrak{G}$ zu einer Untergruppe des direkten Produktes der durch die $A_i(\mathfrak{G})$ dargestellten abstrakten Gruppen isomorph ist, gilt dann auch $r(\mathfrak{G}) \leq n$.

Weiterhin sei $\mathfrak{G}$ absolut irreduzibel.

b) Nun sei $\mathfrak{G}$ imprimitiv mit k Imprimitivitätsgebieten $\mathfrak{r}_i$ der Dimension m. Wir betrachten statt $\mathfrak{G}$ die in Hilfssatz 4 konstruierte Obergruppe $\mathfrak{G}^*$. Wenn wir $r(\mathfrak{G}^*) \leq n$ nachweisen können, so ist auch $r(\mathfrak{G}) \leq n$ gesichert.

Da wir nach a) $\mathfrak{G}^*$ als absolut irreduzibel annehmen können, ist auch $B(\mathfrak{G}_1^*)$ absolut irreduzibel (Satz 2). Auf die Gruppe $B(\mathfrak{G}_1^*)$ vom Grad m können wir nun die Induktionsannahme anwenden. Dies liefert uns $r(B(\mathfrak{G}_1^*)) \leq m$ Damit ist aber die Voraussetzung von Hilfssatz 5 mit m statt c erfüllt, und wir erhalten $r_{\mathfrak{G}^*}(\mathfrak{D}^*) \leq km = n$.

Nun bleibt noch zu zeigen, daß die zu $\mathfrak{G}^*/\mathfrak{D}^*$ isomorphe Permutationsgruppe $\mathfrak{P}_k$ höchstens den Rang n hat. Dazu betrachten wir $\mathfrak{P}_k$ über dem komplexen Zahlkörper. $\mathfrak{P}_k$ enthält die identische Darstellung und ist vollständig reduzibel in absolut irreduzible Teile, die höchstens den Grad $n-1$ haben. Nun liefert die Anwendung der Induktionsannahme auf die Bestandteile von $\mathfrak{P}_k$, daß sogar $r(\mathfrak{P}_k) = r(\mathfrak{G}^*/\mathfrak{D}^*) \leq n-1$ gilt. Damit haben wir in diesem Fall $r(\mathfrak{G}^*) \leq n$ bewiesen.

c) Jetzt lasse sich $\mathfrak{G}$ zerlegen in ein Kronecker-Produkt $\mathfrak{G} = P_1(\mathfrak{G}) \times P_2(\mathfrak{G})$ mit Grad $P_i(\mathfrak{G}) = n_i$. Auf die zu den $P_i(\mathfrak{G})$ gehörigen Gruppen $\mathfrak{R}_i$ können wir unsere Induktionsannahme anwenden; denn wegen der absoluten Irreduzibilität von $\mathfrak{G}$ sind auch die $\mathfrak{R}_i$ absolut irreduzibel. Somit bekommen wir $r(\mathfrak{R}_i) \leq n_i < n$. Nun ist aber $\mathfrak{G}$ ein homomorphes Bild einer Untergruppe des direkten Produktes der $\mathfrak{R}_i$, hat daher ebenfalls höchstens den Rang n.

d) Es ist noch der Fall 3 zu betrachten, daß alle reduziblen Normalteiler von $\mathfrak{G}$ in lauter äquivalente Bestandteile vom Grad 1 zerfallen. Nach Hilfssatz III.2 haben wir dann $|\mathfrak{N}/\mathfrak{Z}(\mathfrak{N})| = n^2 = q^{2m}$. Die Gruppe $\mathfrak{C}(\mathfrak{N}) = \mathfrak{Z}(\mathfrak{G})$ ist nach dem Schurschen Lemma zyklisch, die Faktorgruppe $\mathfrak{C}(\mathfrak{H})/\mathfrak{C}(\mathfrak{N})$ ist nach Hilfssatz II.2 eine abelsche Gruppe mit höchstens $2m$ Erzeugenden. Wegen $2m \leq n = q^m$ haben wir damit $r_{\mathfrak{G}}(\mathfrak{C}(\mathfrak{H})) \leq n$ bewiesen. Es fehlt noch der Nachweis, daß auch $r(\mathfrak{G}/\mathfrak{C}(\mathfrak{H})) = r(\overline{\mathfrak{G}}) \leq n$ ist. $\overline{\mathfrak{G}}$ ist eine über $GF(q)$ irreduzible Untergruppe von $GL(2m, q)$. Nach Hilfssatz 3 zerfällt $\overline{\mathfrak{G}}$ über einem geeigneten Erweiterungskörper von $GF(q)$ vollständig in absolut irreduzible Teile. Also können wir auf $\overline{\mathfrak{G}}$ unsere Induktionsannahme anwenden, sofern $2m < n$ ist. Wir bekommen so $r(\overline{\mathfrak{G}}) \leq 2m < n$ und dann auch $r(\mathfrak{G}) \leq n$.

Nun fehlen uns noch die Sonderfälle $n = 2$ und $n = 4$, in denen die Ungleichung $2m < n$ nicht erfüllt ist.

Für $n = 2$ ist $q = 2$ und daher $\overline{\mathfrak{G}} \leq GL(2, 2)$. Dann ist aber $\overline{\mathfrak{G}}$ sogar überauflösbar und $r(\overline{\mathfrak{G}}) = 1$.

Für $n=4$ kann der Fall 3 von § 2 nicht eintreten, da über $GF(2)$ die Einheitsmatrix E selbst das einzige Vielfache von E ist. Daher liegt für $\overline{\mathfrak{G}}$ einer der schon erledigten Fälle a), b) oder c) vor. Dies zeigt uns $r(\overline{\mathfrak{G}}) \leq 4$ und dann wie oben auch $r(\mathfrak{G}) \leq 4$.

Damit ist der Beweis von Satz 12 beendet.

Der Beweisgang von Satz 12 legt die Vermutung nahe, daß bei geeigneten imprimitiven Gruppen $r(\mathfrak{G})$ wirklich nahe bei n liegt, die angegebene Schranke sich also nicht wesentlich verbessern läßt.

§ 6. Die Ordnung irreduzibler Gruppen

Wir betrachten Gruppen über einem Galois-Feld $GF(p^f)$. Uns interessiert die maximale Potenz der Charakteristik p, welche in der Ordnung einer auflösbaren Gruppe $\mathfrak{G}$ vom Grad n über $GF(p^f)$ aufgehen kann. Die höchste p-Potenz, welche in der Ordnung der vollen linearen Gruppe $GL(n, p^f)$ aufgeht, ist bekanntlich

$$p^{f\frac{n(n-1)}{2}}$$

Diese Potenz kann auch in der Ordnung einer auflösbaren Untergruppe von $GL(n, p^f)$ aufgehen; dazu brauchen wir nur eine p-Sylow-Gruppe von $GL(n, p^f)$ zu nehmen. Setzen wir aber voraus, daß $\mathfrak{G}$ über $GF(p^f)$ vollständig reduzibel ist, so kann die Ordnung von $\mathfrak{G}$ nur wesentlich kleinere Potenzen von p enthalten:

Satz 13. $\mathfrak{G}$ sei auflösbar vom Grad n und vollständig reduzibel über $GF(p^f)$. Ist p^g die höchste p-Potenz, welche die Ordnung von $\mathfrak{G}$ teilt, so gilt für $p>2$ die Abschätzung

$$g \leq f\left(\frac{3n}{p} - 1\right);$$

ist die rechte Seite dieser Ungleichung negativ, so ist die Aussage durch $g=0$ zu ersetzen.

Satz 13 liefert eine wesentlich bessere Abschätzung als

$$g \leq f\,\frac{n(n-1)}{2}.$$

Besonders interessant ist, daß die Schranke mit wachsender Primzahl p besser wird.

Beweis. Wir verwenden wieder das Induktionsverfahren von § 2; jedoch sind erhebliche Verfeinerungen nötig, da sich unsere Aussage jetzt auf einen festen Körper bezieht und wir daher nicht sofort zur algebraisch abgeschlossenen Hülle übergehen können.

a) Zuerst zerfalle $\mathfrak{G}$ über $GF(p^f)$ vollständig in k irreduzible Teile $A_i(\mathfrak{G})$ der Grade n_i. Wenn der Strich bedeutet, daß wir nur über die i mit $3n_i - p > 0$

summieren, so ist nach unserer Induktionsannahme

$$g \leq \sum_{i=1}^{k}{}' f\left(\frac{3}{p} n_i - 1\right) \leq f\left(\frac{3}{p} \sum_{i=1}^{k} n_i - 1\right) \leq f\left(\frac{3n}{p} - 1\right)$$

oder $g = 0$.

Wir können uns daher weiterhin auf die Betrachtung solcher Gruppen $\mathfrak{G}$ beschränken, die über $GF(p^f)$ irreduzibel sind.

b) Nun besitze $\mathfrak{G}$ einen Normalteiler, der über $GF(p^f)$ zerfällt, und zwar in nicht lauter äquivalente irreduzible Bestandteile. Dann ist $\mathfrak{G}$ über $GF(p^f)$ imprimitiv (Satz 2) und enthält einen Normalteiler

$$\mathfrak{D} = \begin{pmatrix} D_1(\mathfrak{D}) & 0 & \dots & 0 \\ 0 & D_2(\mathfrak{D}) & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & D_k(\mathfrak{D}) \end{pmatrix}$$

derart, daß $\mathfrak{G}/\mathfrak{D}$ zu einer Untergruppe $\mathfrak{P}_k$ der symmetrischen Gruppe $\mathfrak{S}_k$ isomorph ist. Da die Teilmoduln $\mathfrak{r}_i$ unter $\mathfrak{G}$ transitiv permutiert werden, haben alle $D_i(\mathfrak{D})$ die gleiche Ordnung. Nach unserer Induktionsannahme gilt nun

$$\text{für } p^e \top |D_i(\mathfrak{D})| \quad \text{die Abschätzung} \quad c \leq f\left(\frac{3}{p}\frac{n}{k} - 1\right)$$

oder $c = 0$,

$$\text{für } p^d \top |\mathfrak{D}| \quad \text{dann} \quad d \leq kc \leq f\left(\frac{3}{p} n - k\right)$$

oder $d = 0$.

Ist schließlich $p^e \top |\mathfrak{G}/\mathfrak{D}|$, so gilt $p^e | k!$. Dann ist bekanntlich

$$e \leq \left[\frac{k}{p}\right] + \left[\frac{k}{p^2}\right] + \dots < \frac{k}{p} + \frac{k}{p^2} + \dots = \frac{k}{p-1}.$$

Insgesamt erhalten wir nun für $d > 0$ die Abschätzung

$$g \leq f\left(\frac{3}{p} n - k\right) + \frac{k}{p-1} \leq f\left(\frac{3}{p} n - k + \frac{k}{p-1}\right).$$

Ist $k\left(1 - \frac{1}{p-1}\right) \geq 1$, so sind wir fertig. Dies ist aber wegen $p > 2$ und $k \geq 2$ sicher der Fall.

Ist aber $d = 0$, so gilt $g = e < \frac{k}{p-1} \leq \frac{n}{p-1}$. Wegen $\frac{p}{2} < p - 1$ gilt daher für $n \geq p$ die gewünschte Ungleichung

$$g < \frac{n}{p-1} < \frac{2n}{p} \leq \frac{3n}{p} - 1.$$

Für $n < p$ ist $p \nmid k!$ und daher $g = e = 0$.

Wir können weiterhin annehmen, daß jeder über $GF(p^l)$ reduzible Normalteiler von $\mathfrak{G}$ in lauter äquivalente irreduzible Bestandteile zerfällt.

c) Nun nehmen wir an, daß es einen abelschen Normalteiler $\mathfrak{A}$ von $\mathfrak{G}$ gibt, der nicht nur aus Matrizen der Gestalt aE besteht. Da $\mathfrak{A}$ nach b) über $GF(p^l)$ in lauter äquivalente irreduzible Teile zerfällt, ist

$$\mathfrak{A}=\begin{pmatrix} D(\mathfrak{A}) & 0 & \dots & 0 \\ 0 & D(\mathfrak{A}) & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & D(\mathfrak{A}) \end{pmatrix}$$

mit einer über $GF(p^l)$ irreduziblen Darstellung $D(\mathfrak{A})$ vom Grad k; dabei ist $1<k\leq n$. Wir setzen $n=jk$, so daß $\mathfrak{A}$ in j Teile zerfällt. Da $D(\mathfrak{A})$ abelsch und über $GF(p^l)$ irreduzibel ist, ist es nach Hilfssatz 2 zyklisch. Folglich läßt sich $D(\mathfrak{A})$ in $GF(p^{lk})$ auf Diagonalgestalt bringen, und dann $\mathfrak{A}$ ebenso. Ist A ein erzeugendes Element von $\mathfrak{A}$, so gilt also in geeigneten Koordinaten

$$A=\begin{pmatrix} a_1 & 0 & \dots & 0 & 0 & \dots & 0 \\ 0 & a_2 & \dots & 0 & 0 & \dots & 0 \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & a_k & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 & a_1 & \dots & 0 \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & 0 & 0 & \dots & a_k \end{pmatrix}$$

mit $a_i\in GF(p^{lk})$. Nochmalige Koordinatenänderung führt zu

$$\begin{pmatrix} a_1E_j & 0 & \dots & 0 \\ 0 & a_2E_j & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & a_kE_j \end{pmatrix},$$

wobei E_j die j-reihige Einheitsmatrix ist. Die a_i sind als Eigenwerte der über $GF(p^l)$ irreduziblen Matrix $D(A)$ alle verschieden (Hilfssatz 2). Daher ist $\mathfrak{G}$ über $GF(p^{lk})$ imprimitiv und hat einen Normalteiler

$$\mathfrak{D}=\begin{pmatrix} D_1(\mathfrak{D}) & 0 & \dots & 0 \\ 0 & D_2(\mathfrak{D}) & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & D_k(\mathfrak{D}) \end{pmatrix}$$

mit $\mathfrak{G}/\mathfrak{D}\simeq\mathfrak{P}_k\leq\mathfrak{S}_k$.

$D_1(\mathfrak{D})$ stellt $\mathfrak{D}$ treu dar: Denn ist $\mathfrak{N}$ der Kern des Homomorphismus von $\mathfrak{D}$ auf $D_1(\mathfrak{D})$, so enthält $\mathfrak{N}$ die identische Darstellung. $\mathfrak{D}$ zerfällt über $GF(p^l)$ höchstens in lauter äquivalente Teile $C(\mathfrak{D})$. Dann enthält $C(\mathfrak{N})$ ebenfalls die identische Darstellung, und nur diese, da wegen der Irreduzibilität von $C(\mathfrak{D})$

über $GF(p^f)$ alle Bestandteile von $C(\mathfrak{N})$ konjugiert sind[5]) (CLIFFORD [3]). Nun folgt $C(\mathfrak{N}) = \mathfrak{E}$ und dann $\mathfrak{N} = \mathfrak{E}$, was zu zeigen war.

Bei der separablen Erweiterung von $GF(p^f)$ zu $GF(p^{fk})$ bleibt $\mathfrak{D}$ vollständig reduzibel (v. d. WAERDEN [18], S. 165). Unsere Induktionsannahme liefert uns daher wegen $|D_1(\mathfrak{D})| = |\mathfrak{D}|$ für $p^d \top |\mathfrak{D}|$ die Abschätzung

$$d \leq fk\left(\frac{3}{p}\,\mathrm{Grad}\,D_1(\mathfrak{D}) - 1\right) = fk\left(\frac{3}{p}\,\frac{n}{k} - 1\right) = f\left(\frac{3n}{p} - k\right)$$

oder $d = 0$.

Ist $p^e \top |\mathfrak{G}/\mathfrak{D}|$, so gilt wieder $p^e | k!$ und dann $e < \frac{k}{p-1}$. Für $d > 0$ erhalten wir nun

$$g \leq f\left(\frac{3n}{p} - k + \frac{k}{p-1}\right).$$

Wegen $k \geqq 2$ können wir nun wie unter b) weiterschließen. Für $d = 0$ bleibt der Schluß von b) unverändert richtig.

d) Es ist noch der Fall zu behandeln, daß alle abelschen Normalteiler von $\mathfrak{G}$ nur aus Vielfachen der Einheitsmatrix E bestehen. Dann betrachten wir wie in § 2 einen minimalen nichtabelschen Normalteiler $\mathfrak{N}$ von $\mathfrak{G}$. Nun liegen die Voraussetzungen der Hilfssätze I bis III vor, wobei $\mathfrak{A} = \mathfrak{Z}(\mathfrak{G})$ nur aus Vielfachen der Einheitsmatrix E besteht. Nach Hilfssatz I.3 ist dann $\mathfrak{C}(\mathfrak{Z}(\mathfrak{N})) = \mathfrak{G}$. Wegen Hilfssatz I.6 gilt $q | p^f - 1$. Gemäß Hilfssatz III.1 lassen sich daher die absolut irreduziblen Darstellungen von $\mathfrak{N}$ für $q \neq 2$ schon in $GF(p^f)$ angeben; für $q = 2$ ist dies auch noch richtig, sofern $p^f \equiv 1(4)$ gilt.

Wir behandeln zunächst den Fall, daß $\mathfrak{N}$ über $GF(p^f)$ in $k > 1$ äquivalente, absolut irreduzible Teile zerfällt. Für $q \neq 2$ und $q = 2$ mit $4 | p^f - 1$ ist dies nach der eben gemachten Feststellung sicher der Fall, falls nicht $\mathfrak{N}$ überhaupt absolut irreduzibel ist. Nun läßt sich $\mathfrak{G}$ nach Satz 3 zerlegen in ein Kronecker-Produkt $\mathfrak{G} = P_1(\mathfrak{G}) \times P_2(\mathfrak{G})$ mit Grad $P_1(\mathfrak{G}) = k$. Die $P_i(\mathfrak{G})$ entstehen aus Gruppen $\mathfrak{N}_i$ des gleichen Grades. Wegen der Irreduzibilität von $\mathfrak{G}$ sind auch die $\mathfrak{N}_i$ irreduzibel über $GF(p^f)$. $\mathfrak{G}$ ist ein homomorphes Bild einer Untergruppe des direkten Produktes der $\mathfrak{N}_i$. Unsere Induktionsannahme liefert uns daher mit $p^{g_i} \top |\mathfrak{N}_i|$

$$\begin{aligned}
&g = 0 && \text{für } g_1 = g_2 = 0,\\
&g \leq f\left(\frac{3k}{p} - 1\right) && \text{für } g_1 \neq 0,\ g_2 = 0,\\
&g \leq f\left(\frac{3}{p}\,\frac{n}{k} - 1\right) && \text{für } g_1 = 0,\ g_2 \neq 0,\\
&g \leq f\left(\frac{3k}{p} - 1\right) + f\left(\frac{3}{p}\,\frac{n}{k} - 1\right)\\
&\quad = f\left(\frac{3}{p}\left(k + \frac{n}{k}\right) - 2\right) && \text{für } g_i \neq 0\ (i = 1, 2).
\end{aligned}$$

In den ersten drei Fällen ist offenbar unsere Behauptung erfüllt; im vierten Fall haben wir noch $1 < k < n$ und $k | n$ zu beachten, womit dann wegen $k + \frac{n}{k} \leq n$ ebenfalls die Behauptung folgt.

[5]) d.h. sie haben die Gestalt $C(GNG^{-1})$ mit geeignetem $G \in \mathfrak{D}$.

Jetzt betrachten wir den Fall, daß $\mathfrak{N}$ über $GF(p^f)$ in k gleiche Teile zerfällt ($k>1$), diese aber nicht absolut irreduzibel sind. Dies kann nur für $q=2$ und $4 \nmid p^f-1$ eintreten. Da $GF(p^{2f})$ primitive 4-te Einheitswurzeln enthält, läßt sich nach Hilfssatz III.1 jede absolut irreduzible Darstellung von $\mathfrak{N}$ jedenfalls in $GF(p^{2f})$ angeben. $\mathfrak{Z}(\mathfrak{N})$ ist ein abelscher Normalteiler von $\mathfrak{G}$, besteht also nur aus Vielfachen der Einheitsmatrix E. Wegen $4 \nmid p^f-1$ ist daher $|\mathfrak{Z}(\mathfrak{N})|=2$. Nach Hilfssatz III.4 besitzt $\mathfrak{N}$ über $GF(p^{2f})$ genau eine absolut irreduzible treue Darstellung.

Nun möge jeder der k über $GF(p^f)$ irreduziblen Teile von $\mathfrak{N}$ beim Übergang zu $GF(p^{2f})$ in i absolut irreduzible Teile zerfallen. Diese Teile sind nach dem eben Gezeigten alle gleich. $\mathfrak{N}$ zerfällt dann also über $GF(p^{2f})$ in ik absolut irreduzible gleiche Teile. Wieder können wir $\mathfrak{G}$ in ein Kronecker-Produkt zerlegen. Den trivialen Fall, daß nicht in den Ordnungen beider Gruppen $\mathfrak{N}_i$ die Primzahl p aufgeht, lassen wir beiseite. Unsere Induktionsannahme liefert uns nun

$$g \leq 2f\left(\frac{3}{p} ik + \frac{3}{p}\frac{n}{ik} - 2\right).$$

Ist $n \geq 2^2 ik$, so haben wir wegen $k \geq 2$ und $i \geq 2$ die Ungleichungen $4 \leq ik \leq \frac{n}{4}$ und daher

$$g \leq 2f\left(\frac{3}{p}\frac{n}{4} + \frac{3}{p}\frac{n}{4} - 2\right) = f\left(\frac{3n}{p} - 4\right) < f\left(\frac{3n}{p} - 1\right),$$

was zu zeigen war.

Übrig bleibt jetzt noch der Fall, daß der Grad n/ik des einzigen vorkommenden absolut irreduziblen Bestandteiles von $\mathfrak{N}$ kleiner als 2^2 ist. Da dieser Grad eine Potenz von 2 ist (Hilfssatz III.3), muß dann $n/ik=2$ sein. Nun haben wir $|\mathfrak{N}/\mathfrak{Z}(\mathfrak{N})| = 2^2$ (Hilfssatz III.3), also wegen $|\mathfrak{Z}(\mathfrak{N})|=2$ dann $|\mathfrak{N}| = 2^3$. $\mathfrak{N}/\mathfrak{Z}(\mathfrak{N})$ ist ein minimaler Normalteiler von $\mathfrak{G}/\mathfrak{Z}(\mathfrak{N})$; dann kann aber $\mathfrak{N}$ keine Dieder-Gruppe sein, ist daher eine Quaternionengruppe $\mathfrak{Q}$; denn die Dieder-Gruppe der Ordnung 8 besitzt eine charakteristische Untergruppe vom Index 2. Die einzige absolut irreduzible, treue Darstellung von $\mathfrak{Q}$ über einem Körper K mit Char. $K \neq 2$ wird geliefert durch die Matrizen

$$A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \quad \text{und} \quad B = \begin{pmatrix} i & j \\ j & -i \end{pmatrix}$$

mit $i^2+j^2=-1$. Da solche i, j in jedem Primkörper $GF(p)$ mit $p \neq 2$ gefunden werden können (SCHOLZ-SCHOENEBERG [13], S. 67), lassen sich alle absolut irreduziblen, treuen Darstellungen von $\mathfrak{Q}$ bereits in $GF(p^f)$ angeben. Damit liegt der zu Beginn von d) schon erledigte Fall vor.

Im folgenden dürfen wir also annehmen, daß der Normalteiler $\mathfrak{N}$ von $\mathfrak{G}$ über $GF(p^f)$ irreduzibel ist.

e) Nun können wir uns auf die Betrachtung der mit $\mathfrak{N}$ ebenfalls über $GF(p^f)$ irreduziblen Gruppe $\mathfrak{N}\mathfrak{G}_p$ beschränken. Wir dürfen annehmen, daß auch $\mathfrak{N}\mathfrak{G}_p$ keinen abelschen Normalteiler besitzt, der nicht nur aus Vielfachen der Einheitsmatrix E besteht; andernfalls ließen sich nämlich die früheren Betrachtungen auf $\mathfrak{N}\mathfrak{G}_p$ anwenden.

Sei $\mathfrak{K}$ ein minimaler nichtabelscher Normalteiler von $\mathfrak{N}\mathfrak{G}_p$. Wir können annehmen, daß $\mathfrak{K}$ über $GF(p^l)$ irreduzibel ist, da sonst einer der schon behandelten Fälle vorläge. Daher dürfen wir in der weiteren Betrachtung $\mathfrak{N}\mathfrak{G}_p$ durch $\mathfrak{K}\mathfrak{G}_p$ ersetzen. Führen wir diesen Schritt mehrfach durch, so können wir schließlich annehmen, daß $\mathfrak{N}$ minimaler nichtabelscher Normalteiler von $\mathfrak{N}\mathfrak{G}_p$ ist.

Ist $q>2$ oder $q=2$ und $4\mid p^l-1$, so ist nach Hilfssatz III.3 $\mathfrak{N}$ absolut irreduzibel und $|\mathfrak{N}/\mathfrak{Z}(\mathfrak{N})|=n^2=q^{2m}$, also $n=q^m$. Ist aber $q=2$ und $4\nmid p^l-1$, also $|\mathfrak{Z}(\mathfrak{N})|=2$, so können wir auf den eindeutig bestimmten absolut irreduziblen Bestandteil von $\mathfrak{N}$ Hilfssatz III.4 anwenden und erhalten, wenn $\mathfrak{N}$ in t absolut irreduzible Teile zerfällt, $n=t2^m$ mit $|\mathfrak{N}/\mathfrak{Z}(\mathfrak{N})|=2^{2m}$.

Wir betrachten die Automorphismen, welche eine p-Sylow-Gruppe von $\mathfrak{G}$ bei Transformation auf $\mathfrak{H}=\mathfrak{N}/\mathfrak{Z}(\mathfrak{N})$ bewirkt. Dabei erfährt $\mathfrak{G}_p$ eine Darstellung $H(\mathfrak{G}_p)$ vom Grad $2m$ über $GF(q)$. $H(\mathfrak{G}_p)$ ist irreduzibel über $GF(q)$; denn nach Hilfssatz I.2 ist $\mathfrak{H}$ Hauptfaktor von $\mathfrak{G}=\mathfrak{N}\mathfrak{G}_p$. Wegen $\mathfrak{C}(\mathfrak{N})=\mathfrak{Z}(\mathfrak{G})$ [Hilfssatz II.1 mit $\mathfrak{Z}(\mathfrak{G})$ statt $\mathfrak{A}$] und der Kommutativität von $\mathfrak{C}(\mathfrak{H})/\mathfrak{C}(\mathfrak{N})$ [Hilfssatz II.2 mit $\mathfrak{C}(\mathfrak{Z}(\mathfrak{N}))=\mathfrak{G}$] ist $\mathfrak{C}(\mathfrak{H})$ nilpotent, kann daher wegen Hilfssatz 1 keine nichttriviale p-Sylow-Gruppe besitzen. Folglich gilt $\mathfrak{C}(\mathfrak{H})\cap\mathfrak{G}_p=\mathfrak{E}$, und $H(\mathfrak{G}_p)$ ist treu.

Zunächst behandeln wir den Fall, daß die p-Sylow-Gruppe $\mathfrak{G}_p$ abelsch ist. Da $\mathfrak{G}_p$ durch $H(\mathfrak{G}_p)$ treu und irreduzibel dargestellt wird, ist $\mathfrak{G}_p$ nach Hilfssatz 2 zyklisch. Sei P ein erzeugendes Element von $\mathfrak{G}_p$. Bekanntlich läßt sich P über $GF(p^l)$ auf Dreiecksgestalt bringen (Speiser [15], S. 222). Daher ist für $p^a\geq n$

$$P^{p^a}=E.$$

Wegen $|\mathfrak{G}_p|=p^g$ gilt also $p^{g-1}<n$.

Sei zunächst $n<p$. Da $H(\mathfrak{G}_p)$ über $GF(q)$ irreduzibel ist, kann es in $GF(q^{2m})$ auf Diagonalgestalt gebracht werden. Daher haben wir

$$p^g\mid q^{2m}-1=(q^m-1)(q^m+1).$$

Wegen $p>n\geq q^m$ folgt $p^g\mid q^m+1$. Abgesehen von dem Fall $g=0$ ist dies nur für $g=1$ und $p=n+1=q^m+1$ erfüllt. Dann folgt aber die Richtigkeit unserer Behauptung unter Beachtung von $p\geq 3$ aus

$$\frac{3n}{p}-1=\frac{3(p-1)}{p}-1=\frac{2p-3}{p}\geq 1\geq g.$$

Nun sei $n\geq p$. Angenommen, es sei entgegen unserer Behauptung $g>f\left(\frac{3n}{p}-1\right)\geq\frac{3n}{p}-1$. Dann bekommen wir wegen $\frac{3n}{p}-2\geq 1$ mit Hilfe der Ungleichung von Bernoulli

$$n>p^{g-1}>p^{3n/p-2}=(1+(p-1))^{3n/p-2}\geq 1+(p-1)\left(\frac{3n}{p}-2\right)$$

und dann

$$\frac{n(3-2p)}{p}>3-2p,$$

wegen $3-2p<0$ also $n<p$, entgegen unserer Annahme $n \geq p$. Damit ist für abelsches $\mathfrak{G}_p$ unser Satz bewiesen.

f) Es bleibt noch der Fall $\mathfrak{G}_p' \neq \mathfrak{E}$ zu behandeln. Wir betrachten die absolut irreduziblen Bestandteile der über $GF(q)$ irreduziblen Darstellung $H(\mathfrak{G}_p)$. Diese sind nichtlinear, haben daher alle denselben Grad (Hilfssatz 3), welcher wegen $q \neq p$ durch p teilbar ist. Also gilt $2m = pj$ mit einer geeigneten ganzen Zahl j.

Nach Hilfssatz II.3 ist $H(\mathfrak{G}_p)$ eine Untergruppe der symplektischen Gruppe $C(2m, q)$. Daraus folgt nun

$$p^g \mid |C(2m,q)| = (q^{2m}-1)\, q^{2m-1}(q^{2m-2}-1)\, q^{2m-3} \ldots q^3(q^2-1)\, q$$

und wegen $q \neq p$ dann

$$p^g \mid (q^{2m}-1)\,(q^{2m-2}-1) \ldots (q^2-1).$$

(Die Ordnung von $C(2m, q)$ entnehmen wir aus v. d. WAERDEN [17], S. 10.)

Zwischen p und q bestehen nun mehrere Bindungen:

Zunächst ist $q \mid p^j - 1$, also

$$q < p^j. \tag{1}$$

Sei r die Ordnung von q mod p, also

$$q^r = 1 + y p^x \quad \text{mit} \quad p \nmid y \text{ und } x \geq 1. \tag{2}$$

Dann ist $r \mid q(p) = p-1$, daher

$$r \leq p-1. \tag{3}$$

Aus (2) folgt $p^x < q^r$, dann mit (1) $p^{x/r} < q < p^j$, also $x/r < j$ und daher

$$x \leq rj - 1. \tag{4}$$

Nun wollen wir die maximale p-Potenz, welche in $|C(2m, q)|$ aufgeht, abschätzen. Dabei bedienen wir uns einer Rechnung, die BURNSIDE im Spezialfall $f = 1$ bei einer ähnlichen Gelegenheit durchgeführt hat (BURNSIDE [2]).

Aus $q^r = 1 + y p^x$ folgt für $v = rkp^j$ mit $p \nmid k$

$$q^v = (1 + y p^x)^{k p^j} = 1 + k y p^{x+j} + \binom{k p^j}{2} y^2 p^{2x} + \cdots.$$

Wegen $p \neq 2$ ist also $p^{x+j} \parallel q^v - 1$.

Sei c die ganze Zahl mit

$$cr \leq 2m < (c+1)r. \tag{5}$$

Wir bestimmen jetzt die höchste p-Potenz p^u, welche in

$$(q^{2m}-1)\,(q^{2m-1}-1) \ldots (q^2-1)\,(q-1)$$

aufgeht. Durch p teilbar sind genau die c Faktoren $q^{ri}-1$ mit $i \leq c$, und jeder von diesen liefert den Beitrag p^x. Je einen weiteren Faktor p erhalten

wir aus den $[c/p]$ Gliedern mit $p|i$, noch je einen weiteren Faktor p aus den $[c/p^2]$ Gliedern mit $p^2|i$ usw. Schließlich bekommen wir so

$$u = c\,x + \left[\frac{c}{p}\right] + \left[\frac{c}{p^2}\right] + \cdots .$$

Ist $c = a_0 + a_1 p + a_2 p^2 + \cdots + a_s p^s$, so wird

$$u = c\,x + a_1 + a_2(1+p) + \cdots + a_s(1+p+\cdots+p^{s-1}) = c\,x + \frac{c - a_0 - a_1 - \cdots - a_s}{p-1}$$

und daher

$$u \leq c\,x + \frac{c-1}{p-1}. \tag{6}$$

Nun sei zunächst $r = 1$. Dann ist nach (4) $x \leq f-1$ und nach (5) $c \leq 2m$. Wir nehmen an, daß unsere Behauptung nicht stimmt, also $g > f\left(\frac{3n}{p} - 1\right)$ ist und bekommen mit (6)

$$f\left(\frac{3n}{p} - 1\right) < g \leq u \leq 2m(f-1) + \frac{2m-1}{p-1}.$$

Mit $n \geq q^m$ und $2m = jp$ wird daher

$$f\left(\frac{3}{p}\, q^{jp/2} - 1\right) < jp(f-1) + \frac{jp-1}{p-1}.$$

Multiplikation mit p liefert

$$f(3q^{jp/2} - p) < jp^2(f-1) + \frac{p(jp-1)}{p-1} < jp^2\left(f - 1 + \frac{1}{p-1}\right) < jp^2 f$$

oder

$$3q^{jp/2} < jp^2 + p = p(jp+1). \tag{I}$$

Erst recht ist dann

$$3q^{jp/2} < 2j^2p^2, \quad \text{also mit} \quad jp = 2m = z \qquad 3q^{z/2} < 2z^2.$$

Dabei interessieren uns nur die geraden z. Im ungünstigsten Falle $q = 2$ bekommen wir nun $3\,2^{z/2} < 2z^2$. Eine einfache Rechnung zeigt, daß dies für $z \geq 16$ nicht stimmt. Es ist leicht zu sehen, daß die restlichen j, p mit $z = jp \leq 14$, geradem j und $p \neq 2$ die Ungleichung (I) nicht erfüllen. Damit sind wir zu einem Widerspruch gekommen, wodurch im Falle $r = 1$ unsere Behauptung bewiesen ist.

Nun sei $r > 1$. Nach (3) und (4) ist $x \leq rf - 1 \leq (p-1)f - 1$, nach (5) $c \leq m$. Wieder nehmen wir an, es sei entgegen unserer Behauptung $g > f\left(\frac{3n}{p} - 1\right)$. Dann wird nach (6)

$$f\left(\frac{3n}{p} - 1\right) < g \leq u \leq m\left((p-1)f - 1\right) + \frac{m-1}{p-1}.$$

Unter Beachtung von $n \geq q^m$ und $2m = jp$ folgt

$$f\left(\frac{3}{p}\, q^{jp/2} - 1\right) < \frac{jp}{2}\left((p-1)f - 1\right) + \frac{jp/2 - 1}{p-1} < \frac{jp}{2}(p-1)f,$$

daher

$$\binom{3}{p} q^{jp/2} < \binom{jp}{2} (p-1) + 1.$$

Nun wird

(II) $$3q^{jp/2} < \binom{jp^2}{2} (p-1) + p.$$

Daraus ergibt sich wegen $q \geq 2$

$$3\,2^{jp/2} < \frac{jp^3}{2} - \frac{jp^2}{2} + p < \frac{jp^3}{2} < \frac{j^3p^3}{2}.$$

Mit $jp = 2m = z$ erhalten wir die Ungleichung $3\,2^{z/2+1} < z^3$, welche für $z > 20$ nicht erfüllt ist. Nun haben wir noch nachzuprüfen, welche Paare j, p mit $jp \leq 20$ die Ungleichung (II) erfüllen; dabei fallen wegen $p \neq 2$ noch die reinen 2-Potenzen weg. Es ergibt sich, daß (II) nur für $p = 5$, $j = 2$ mit $q = 2$ erfüllt ist. Die genaue 5-Potenz, welche in $(2^{10}-1)(2^9-1)\ldots(2^2-1)$ aufgeht, ist aber 5^2. Daher gilt in diesem Falle wegen $n \geq 2^5$ die Ungleichung $g \leq 2 < f\left(\frac{3}{5}\,2^5 - 1\right) \leq f\left(\frac{3n}{5} - 1\right)$, was zu zeigen war.

Damit haben wir den Beweis von Satz 13 beendet.

Für $p = 2$ gilt ein ähnlicher Satz:

Satz 14. Sei $\mathfrak{G}$ auflösbar und vollständig reduzibel vom Grad n über $GF(2^f)$. Ist 2^g die höchste 2-Potenz, welche die Ordnung von $\mathfrak{G}$ teilt, so ist $g \leq f(n-1)$.

Beweis. Der gruppentheoretische Teil des Beweises verläuft wie bei Satz 13. Wir wollen uns daher kurz fassen und geben nur die Abweichungen bei den Abschätzungen an.

a) Der Schluß verläuft ebenso wie bei Satz 13 mit $f(n_i - 1)$ an der Stelle von $f(3n_i/p - 1)$.

b) Nun wird $d \leq f(n-k)$ und $e < k$, also $e \leq k-1$ und dann

$$g \leq f(n-k) + k - 1 \leq f(n-k+k-1) = f(n-1).$$

c) Jetzt erhalten wir $d \leq f(n-k)$ und wieder $e \leq k-1$, also wie eben $g \leq f(n-1)$.

d) Statt S. 505 unten muß es jetzt heißen:

$$g \leq f\left(k - 1 + \binom{n}{k} - 1\right) < f(n-1).$$

Der Fall, daß $\mathfrak{R}$ über $GF(2^f)$ nicht in absolut irreduzible Teile zerfällt, kann wegen $q \neq 2$ nicht eintreten. Daher entfällt jetzt die Betrachtung von S. 506.

e) Der Fall $n = 1$ ist trivial. Angenommen, es sei $g > f(n-1)$. Dann erhalten wir wegen $n > 1$

$$n > 2^{g-1} > 2^{f(n-1)-1}.$$

Daraus folgt $f(n-1) - 1 < n$ und daher $f \leq 3$.

Für $f=1$ bekommen wir $n>2^{n-2}$, also $n=2$ oder $n=3$. Im Falle $n=2$ ist aber wegen $2 \top |GL(2,2)|$ sicher $g \leq 1$, daher unsere Behauptung richtig. Im Falle $n=3$ ist zwar $2^3 \top |GL(3,2)|$, aber die Gruppe $GL(3,2)$ ist nicht auflösbar. Die auflösbaren Untergruppen von $GL(3,2)$, welche die volle 2-Sylow-Gruppe von $GL(3,2)$ umfassen, sind isomorph zur symmetrischen Gruppe $\mathfrak{S}_4$ oder zu der 2-Sylow-Gruppe selbst. Diese Gruppen sind aber nicht vollständig reduzibel, da sie Normalteiler der Ordnung 4 besitzen. Also gibt es keine auflösbare, vollständig reduzible Untergruppe $\mathfrak{G}$ von $GL(3,2)$ mit $2^3 \top |\mathfrak{G}|$. Daher ist $g \leq 2$ und unsere Behauptung erfüllt.

Im Falle $f=2$ erhalten wir $n>2^{2n-3}$ und daher $n=1$; für $f=3$ ergibt sich $n>2^{3n-4}$ und dann ebenfalls $n=1$.

f) Nun haben wir $r=1$, also

$$x \leq f-1 \tag{4'}$$

und

$$c \leq 2m < c+1, \tag{5'}$$

also $c=2m$.

Sei zunächst $q=1+y\,2^x$ mit $2 \nmid y$ und $x>1$. Dann ist mit $v=k\,2^j$ und $2 \nmid k$

$$q^v \equiv 1 + k\,y\,2^{x+j} \bmod 2^{j-1+2x}.$$

Wegen $j-1+2x>x+j$ ist nun wieder $2^{x+j} \top q^v-1$. Unsere frühere Bestimmung von u bleibt daher richtig und liefert

$$u \leq c\,x + c - 1. \tag{6'}$$

Wir nehmen nun wieder an, daß unsere Behauptung nicht richtig ist, also $g>f(n-1)$ gilt. Mit (4'), (5') und (6') ergibt das

$$f(n-1) < g \leq u \leq c\,x + c - 1 \leq 2m(f-1) + 2m - 1 < 2mf.$$

Unter Beachtung von $n \geq q^m \geq 3^m$ folgt nun $3^m < 2m+1$. Diese Ungleichung ist aber für kein $m \geq 0$ erfüllt. Damit ist der gewünschte Widerspruch erreicht.

Endlich bleibt noch der Fall $x=1$ zu behandeln. (BURNSIDE hatte in [2] diesen Fall übersehen; jedoch läßt sich die Lücke bei ihm unschwer schließen.) Nun ist $r=1$ und $q=1+2y$ mit $2 \nmid y$. Wieder haben wir $n \geq q^m \geq 3^m$.

Es ist $q^2-1=4y(y+1)$, also mit $y+1=a\,2^w$ und $2 \nmid a$ dann $2^{w+2} \top q^2-1$. Für $v=k\,2^j$ mit $2 \nmid k$ und $j \geq 1$ erhalten wir daher

$$q^v \equiv 1 + a\,k\,y\,2^{w+j+1} \bmod 2^{2w+j+2},$$

also $2^{w+j+1} \top q^v-1$.

Für die maximale 2-Potenz 2^b, welche in

$$(q^{2^m}-1)(q^{2^{m-2}}-1)\ldots(q^2-1)$$

aufgeht, bekommen wir, entsprechend den früheren Überlegungen

$$b = m(w+2) + \left[\frac{m}{2}\right] + \left[\frac{m}{2^2}\right] + \cdots < m(w+2) + m = m(w+3).$$

Nun nehmen wir an, es sei entgegen unserer Behauptung $g > f(n-1)$. Dann wird

$$f(3^m - 1) \leq f(n-1) < g \leq b \leq m(w+3).$$

Aus $q \mid 2^f - 1$ erhalten wir $q < 2^f$. Andererseits ist $q = a\,2^{w+1} - 1 \geq 2^w$, daher also $w < f$. Es folgt $w \leq f-1$ und $f(3^m-1) < m(f+2) \leq 2mf$, also $3^m < 2m+1$. Die letzte Ungleichung ist aber für kein $m \geq 0$ erfüllt. Damit haben wir die Annahme $g > f(n-1)$ zum Widerspruch geführt und den Beweis von Satz 14 abgeschlossen.

Aus Satz 13 folgt, daß für $n \leq p/3$ die Ordnung von $\mathfrak{G}$ zu p teilerfremd ist. Diese Aussage wollen wir nun noch verschärfen:

Satz 15. Sei $\mathfrak{G}$ auflösbar und vollständig reduzibel vom Grad n über $GF(p^f)$. Ist $n \leq p-1$, so ist die Ordnung von $\mathfrak{G}$ zu p teilerfremd, ausgenommen den Fall, daß $n = p-1$ und p eine Fermatsche Primzahl ist.

Beweis. Wir gehen die Stufen des Beweises von Satz 13 durch und geben die Änderungen an:

a) Da die irreduziblen Bestandteile p-freie Ordnungen haben, gilt dies auch für $\mathfrak{G}$.

b) Nun ist $d = 0$ und $k < p$, daher $e = 0$ und dann $g = 0$.

c) Aus $d = e = 0$ folgt wieder $g = 0$.

d) Mit $g_1 = g_2 = 0$ ergibt sich $g = 0$, und zwar in allen unter d) betrachteten Fällen.

e) Wie auf S. 507 erhalten wir aus $n \leq p-1$ entweder $g = 0$ oder $p = n+1 = q^m + 1$. Dann ist notwendig $q = 2$ und p eine Fermatsche Primzahl.

f) Wegen $n \geq q^m$ und $p \mid 2m$ bekommen wir $2^m \leq q^m \leq n \leq p-1 \leq 2m-1$. Diese Ungleichung ist jedoch für kein $m \geq 0$ erfüllt, so daß dieser Fall unter der Voraussetzung $n \leq p-1$ nicht eintreten kann.

Damit haben wir alle möglichen Fälle betrachtet und Satz 15 bewiesen.

Es sei noch bemerkt, daß ein direkter Beweis von Satz 15 gegenüber dem Beweis von Satz 13 wesentliche Vereinfachungen gestattet; denn da in Satz 15 der Parameter f nicht in der Aussage vorkommt, können wir sofort zu einem genügend großen Galois-Feld übergehen, über welchem alle zu betrachtenden Normalteiler in absolut irreduzible Bestandteile zerfallen.

Schließlich vermerken wir noch, daß Satz 15 wenigstens für Primzahlen p, die keine Fermat-Zahlen sind, scharf ist. Denn es gibt stets irreduzible monomiale Gruppen vom Grad p, deren Ordnung dann notwendig durch p teilbar ist.

§ 7. Hauptreihen auflösbarer Gruppen

In diesem Paragraphen sei $\mathfrak{G}$ stets endlich.

Auf jedem Hauptfaktor $\mathfrak{H}$ der Ordnung p^n von $\mathfrak{G}$ bewirkt $\mathfrak{G}$ bei Transformation eine Automorphismengruppe $\mathfrak{G}(\mathfrak{H})$, welche zu einer auflösbaren, über $GF(p)$ irreduziblen Untergruppe von $GL(n, p)$ isomorph ist. Uns interessiert der Durchschnitt $\mathfrak{F}(p)$ aller Kerne der homomorphen Abbildungen

$\mathfrak{G} \to \mathfrak{G}(\mathfrak{H})$, wenn $\mathfrak{H}$ alle Hauptfaktoren von $\mathfrak{G}$ von p-Potenzordnung (kurz: p-Hauptfaktoren) durchläuft. Unsere späteren Aussagen werden vielfach die Faktorgruppe $\mathfrak{G}/\mathfrak{F}(p)$ betreffen.

Eine Gruppe $\mathfrak{G}$ heißt bekanntlich p-nilpotent, wenn sie ein invariantes p-Sylow-Komplement ${}_p\mathfrak{G}$ besitzt.

Hilfssatz 6. Sei $\mathfrak{G}$ eine p-auflösbare Gruppe. Dann ist der Zentralisator aller p-Hauptfaktoren von $\mathfrak{G}$ der maximale p-nilpotente Normalteiler $\mathfrak{F}(p)$ von $\mathfrak{G}$. Ist $\mathfrak{G}$ auflösbar, so ist der maximale nilpotente Normalteiler $\mathfrak{F}$ der Zentralisator aller Hauptfaktoren.

Beweis. Offenbar ist die Vereinigung p-nilpotenter Normalteiler $\mathfrak{N}_i$ $(i = 1, 2)$ von $\mathfrak{G}$ wieder ein solcher; denn die Vereinigung der (als charakteristische Untergruppen der $\mathfrak{N}_i$) in $\mathfrak{G}$ invarianten p-Sylow-Komplemente ${}_p\mathfrak{N}_i$ der $\mathfrak{N}_i$ ergibt ein invariantes p-Sylow-Komplement von $\mathfrak{N}_1\mathfrak{N}_2$.

a) Sei $\mathfrak{N}$ ein p-nilpotenter Normalteiler von $\mathfrak{G}$. Wir haben zu zeigen, daß $\mathfrak{N}$ alle p-Hauptfaktoren von $\mathfrak{G}$ zentralisiert. Zu diesem Zweck bilden wir die Faktorgruppe $\overline{\mathfrak{G}} = \mathfrak{G}/{}_p\mathfrak{N}$ von $\mathfrak{G}$ nach dem p-Sylow-Komplement ${}_p\mathfrak{N}$ von $\mathfrak{G}$. Da ${}_p\mathfrak{N}$ eine zu p teilerfremde Ordnung hat, genügt der Nachweis unserer Behauptung für $\overline{\mathfrak{G}}$. Die Automorphismengruppen auf den p-Hauptfaktoren stellen sich nun dar als lineare irreduzible Gruppen über $GF(p)$, besitzen daher keinen von $\mathfrak{E}$ verschiedenen Normalteiler von p-Potenzordnung (Hilfssatz 1). Daher bewirkt $\mathfrak{N}/{}_p\mathfrak{N}$ auf jedem p-Hauptfaktor den identischen Automorphismus, was zu zeigen war.

b) Nun sei $\mathfrak{N}$ ein Normalteiler von $\mathfrak{G}$, der auf jedem p-Hauptfaktor von $\mathfrak{G}$ den identischen Automorphismus bewirkt. Dann erleiden insbesondere die p-Hauptfaktoren von $\mathfrak{N}$ selbst unter $\mathfrak{N}$ nur den identischen Automorphismus.

Wir zeigen nun, daß eine p-auflösbare Gruppe $\mathfrak{G}$, welche alle ihre p-Hauptfaktoren zentralisiert, p-nilpotent ist. Dazu verwenden wir einen Induktionsschluß nach der Ordnung von $\mathfrak{G}$.

Hat $\mathfrak{G}$ einen Normalteiler $\mathfrak{M}$ von zu p teilerfremder Ordnung, so ist nach unserer Induktionsannahme $\mathfrak{G}/\mathfrak{M}$ p-nilpotent, dann aber auch $\mathfrak{G}$ selbst. Hat aber $\mathfrak{G}$ einen Normalteiler $\mathfrak{L}$ von der Ordnung p^n, so ist $n = 1$, da alle p-Hauptfaktoren von $\mathfrak{G}$ zentralisiert werden. Für $\mathfrak{G}/\mathfrak{L}$ gilt nach der Induktionsannahme unsere Behauptung, so daß ${}_p\mathfrak{G}\mathfrak{L}$ in $\mathfrak{G}$ invariant ist[6]). Da aber $\mathfrak{L}$ im Zentrum von $\mathfrak{G}$ liegt, folgt nun ${}_p\mathfrak{G}\mathfrak{L} = {}_p\mathfrak{G} \times \mathfrak{L} \trianglelefteq \mathfrak{G}$ und dann ${}_p\mathfrak{G} \trianglelefteq \mathfrak{G}$, was zu beweisen war.

Der Durchschnitt aller $\mathfrak{F}(p)$, wenn p die Primteiler von $|\mathfrak{G}|$ durchläuft, ist der maximale nilpotente Normalteiler $\mathfrak{F}$ von $\mathfrak{G}$. Ist $\mathfrak{G}$ auflösbar, so ist also $\mathfrak{F}$ der Zentralisator aller Hauptfaktoren von $\mathfrak{G}$.

Satz 16. Sei $\mathfrak{G}$ p-auflösbar. Genau dann ist $l_p(\mathfrak{G}) = 1$, wenn $\mathfrak{G}$ auf allen p-Hauptfaktoren Automorphismengruppen von zu p teilerfremder Ordnung bewirkt.

Beweis. Ist $l_p(\mathfrak{G}) = 1$, so gibt es eine Normalteilerkette $\mathfrak{E} \trianglelefteq \mathfrak{N}_1 \trianglelefteq \mathfrak{N}_2 \trianglelefteq \mathfrak{G}$ von $\mathfrak{G}$ mit $p \nmid |\mathfrak{N}_1|$, $p \nmid |\mathfrak{G}/\mathfrak{N}_2|$ und $p^s = |\mathfrak{N}_2/\mathfrak{N}_1|$. Dann ist $\mathfrak{G}_p \leq \mathfrak{N}_2 \leq \mathfrak{F}(p)$.

[6]) ${}_p\mathfrak{G}$ existiert, da $\mathfrak{G}$ p-auflösbar ist.

Daher wird jeder p-Hauptfaktor von $\mathfrak{G}$ durch alle p-Sylow-Gruppen $\mathfrak{G}_p$ zentralisiert, erfährt also eine Automorphismengruppe von zu p teilerfremder Ordnung.

Bewirkt umgekehrt $\mathfrak{G}$ auf allen p-Hauptfaktoren Automorphismengruppen von zu p teilerfremder Ordnung, so gilt $\mathfrak{G}_p \leq \mathfrak{F}(p)$. Dann ist $\mathfrak{F}(p)/_p\mathfrak{F}(p) \cong \mathfrak{G}_p$, und $\mathfrak{E} \trianglelefteq_p \mathfrak{F}(p) \trianglelefteq \mathfrak{F}(p) \trianglelefteq \mathfrak{G}$ liefert eine Normalteilerkette von $\mathfrak{G}$, in der alle p-Hauptfaktoren hintereinander vorkommen, so daß $l_p(\mathfrak{G}) = 1$ gilt.

Aus Satz 16 folgen sofort zwei weitere Aussagen:

Satz 17. Ist $\mathfrak{G}$ p-auflösbar und $\mathfrak{G}_p$ abelsch, so gilt $l_p(\mathfrak{G}) = 1$.

Beweis. Der Satz ergibt sich aus Satz 16.

Satz 17 ist enthalten in einer allgemeineren Aussage über $l_p(\mathfrak{G})$, welche P. HALL und G. HIGMAN in [8] angegeben haben. Ist nämlich $c_p(\mathfrak{G})$ die Klasse von $\mathfrak{G}_p$, so gilt $l_p(\mathfrak{G}) \leq c_p(\mathfrak{G})$.

Satz 18. Ist $\mathfrak{G}$ auflösbar und $r_p(\mathfrak{G}) \leq p-1$, so gilt $l_p(\mathfrak{G}) = 1$, ausgenommen den Fall, daß $r_p(\mathfrak{G}) = p-1$ und p eine Fermatsche Primzahl ist.

Beweis. Die Aussage folgt direkt aus den Sätzen 15 und 16.

Die Gruppen vom Rang 1 (die überauflösbaren Gruppen) haben zwei bemerkenswerte Eigenschaften:

Die Sylow-Gruppe zum größten Primteiler der Gruppenordnung ist invariant (P. HALL [7]).

Die Kommutatorgruppe $\mathfrak{G}'$ ist nilpotent (WENDT [19]).

In einer früheren Arbeit (HUPPERT [9]) hatte der Verfasser analoge Sätze für Gruppen vom Rang 2 und 3 angegeben (Satz 14 bis Satz 17 in [9]). Unser systematisches Studium der linearen auflösbaren Gruppen erlaubt jetzt die Aufstellung entsprechender Sätze bei beliebigem Rang.

Wir befassen uns zunächst mit der Verallgemeinerung der ersten Aussage. Dabei wird es sich um eine Abschätzung von $l_p(\mathfrak{G})$ durch den p-Rang $r_p(\mathfrak{G})$ handeln. Allerdings ist zu bemerken, daß dies nur ein schwacher Ersatz für die invariante p-Sylow-Gruppe ist. Denn selbst dann, wenn $l_p(\mathfrak{G}) = 1$ für alle Primteiler p von $|\mathfrak{G}|$ gilt, hat $\mathfrak{G}$ nicht notwendig eine invariante Sylow-Gruppe; dies zeigt das Beispiel $\mathfrak{S}_3 \times \mathfrak{A}_4$ des direkten Produktes der symmetrischen Gruppe $\mathfrak{S}_3$ mit der alternierenden Gruppe $\mathfrak{A}_4$.

Wir zeigen zunächst

Satz 19. Sei $\mathfrak{G}$ auflösbar vom p-Rang $r_p(\mathfrak{G})$. Für $p > 2$ gilt dann

$$r_p\big((\mathfrak{G}/\mathfrak{F}(p)\big) \leq \frac{3}{p} r_p(\mathfrak{G}) - 1,$$

für $p = 2$

$$r_2\big(\mathfrak{G}/\mathfrak{F}(2)\big) \leq r_2(\mathfrak{G}) - 1.$$

Beweis. Sei $\mathfrak{H}$ ein Hauptfaktor von $\mathfrak{G}$ mit $|\mathfrak{H}| = p^n$. Dann ist $n \leq r_p(\mathfrak{G})$. Die Gruppe $\mathfrak{G}(\mathfrak{H})$ ist isomorph zu einer auflösbaren, über $GF(p)$ irreduziblen Untergruppe von $GL(n, p)$. Daher geht für $p > 2$ nach Satz 13 in $|\mathfrak{G}(\mathfrak{H})|$ die

Primzahl p höchstens in der $(3n/p-1)$-ten Potenz auf. Insbesondere ist dann

$$r_p(\mathfrak{G}(\mathfrak{H})) \leq \frac{3n}{p} - 1 \leq \frac{3}{p} r_p(\mathfrak{G}) - 1.$$

Da aber $\mathfrak{G}/\mathfrak{F}(p)$ Untergruppe des direkten Produktes aller der $\mathfrak{G}(\mathfrak{H})$ (bei festgehaltenem p) ist, folgt nun

$$r_p(\mathfrak{G}/\mathfrak{F}(p)) \leq \frac{3}{p} r_p(\mathfrak{G}) - 1.$$

Die Aussage für $p=2$ ergibt sich unter Benutzung von Satz 14 in gleicher Weise.

Jetzt folgt leicht die angekündigte Abschätzung von $l_p(\mathfrak{G})$ durch $r_p(\mathfrak{G})$:

Satz 20. Sei $\mathfrak{G}$ auflösbar vom p-Rang $r_p(\mathfrak{G})$ mit $p \geq 3$. Ist i eine ganze Zahl mit

$$\frac{3^i}{p^i} r_p(\mathfrak{G}) - \left(\frac{3^{i-1}}{p^{i-1}} + \frac{3^{i-2}}{p^{i-2}} + \cdots + \frac{3}{p} + 1 \right) \leq p-2,$$

so gilt $l_p(\mathfrak{G}) \leq i+1$.

Beweis. Nach Satz 19 haben wir $r_p(\mathfrak{G}/\mathfrak{F}(p)) \leq \frac{3}{p} r_p(\mathfrak{G}) - 1$. Die p-Hauptfaktoren von $\mathfrak{G}$ unterhalb $\mathfrak{F}(p) = \mathfrak{F}^1(p)$ lassen sich zu einem p-Abschnitt zusammenfassen, da $\mathfrak{F}(p)$ p-nilpotent ist. Nun betrachten wir denjenigen Normalteiler $\mathfrak{F}^2(p)$ von $\mathfrak{G}$, welcher dem maximalen p-nilpotenten Normalteiler von $\mathfrak{G}/\mathfrak{F}^1(p)$ entspricht. Nochmalige Anwendung von Satz 19 ergibt

$$r_p(\mathfrak{G}/\mathfrak{F}^2(p)) \leq \frac{3}{p} \left(\frac{3}{p} r_p(\mathfrak{G}) - 1 \right) - 1 = \frac{3^2}{p^2} r_p(\mathfrak{G}) - \frac{3}{p} - 1.$$

Entspricht $\mathfrak{F}^i(p)$ dem maximalen p-nilpotenten Normalteiler von $\mathfrak{G}/\mathfrak{F}^{i-1}(p)$, so kommen wir schließlich zu

$$r_p(\mathfrak{G}/\mathfrak{F}^i(p)) \leq \frac{3^i}{p^i} r_p(\mathfrak{G}) - \left(\frac{3^{i-1}}{p^{i-1}} + \frac{3^{i-2}}{p^{i-2}} + \cdots + \frac{3}{p} + 1 \right).$$

Die p-Hauptfaktoren zwischen $\mathfrak{F}^i(p)$ und $\mathfrak{F}^{i+1}(p)$ lassen sich zusammenwerfen, so daß in einer geeigneten Hauptreihe von $\mathfrak{G}$ unterhalb $\mathfrak{F}^i(p)$ höchstens i p-Abschnitte vorkommen. Ist nun i gemäß der obenstehenden Bedingung gewählt, so ist $r(\mathfrak{G}/\mathfrak{F}^i(p)) \leq p-2$ und $\mathfrak{G}/\mathfrak{F}^i(p)$ trägt daher höchstens noch einen weiteren p-Abschnitt bei (s. Satz 18). Damit haben wir $l_p(\mathfrak{G}) \leq i+1$ bewiesen.

Satz 20a. Stets ist $l_p(\mathfrak{G}) \leq r_p(\mathfrak{G})$.

Beweis. Dies ergibt sich nach dem gleichen Verfahren wie Satz 20, wenn wir die aus Satz 19 für alle p folgende Ungleichung $r_p(\mathfrak{G}/\mathfrak{F}(p)) \leq r_p(\mathfrak{G}) - 1$ verwenden.

Wesentlich bessere Abschätzungen erhalten wir für $p>3$:

Satz 20b. Für $p>3$ gilt $l_p(\mathfrak{G}) \leq \left\lfloor p/3 \log \frac{r_p(\mathfrak{G})}{p-1} \right\rfloor + 2$.

Beweis. Die Ungleichung

$$\frac{3^i}{p^i} r_p(\mathfrak{G}) - \left(\frac{3^{i-1}}{p^{i-1}} + \frac{3^{i-2}}{p^{i-2}} + \cdots + \frac{3}{p} + 1\right) \leq p - 2$$

ist sicher dann erfüllt, wenn wir i so wählen, daß

$$\frac{3^i}{p^i} r_p(\mathfrak{G}) - 1 \leq p - 2$$

ist, d.h. $i \geq {}_{p/3}\log \frac{r_p(\mathfrak{G})}{p-1}$. Aus Satz 20 folgt dann die Behauptung.

Aus den Ergebnissen von P. HALL und G. HIGMAN in [8] lassen sich ebenfalls Ungleichungen zwischen $r_p(\mathfrak{G})$ und $l_p(\mathfrak{G})$ herleiten:

Sei $r_p(\mathfrak{G}) = r$. Dann hat jeder p-Hauptfaktor $\mathfrak{H}$ von $\mathfrak{G}$ höchstens r Erzeugende. Eine p-Sylow-Gruppe der auf dem p-Hauptfaktor $\mathfrak{H}$ bewirkten Automorphismengruppe $\mathfrak{G}(\mathfrak{H})$ besteht daher (bezüglich einer geeigneten Basis von $\mathfrak{H}$) aus Dreiecksmatrizen vom Grad $h \leq r$ mit Diagonalelementen 1 (SPEISER [15], S. 222). Bekanntlich hat eine solche Gruppe höchstens die Klasse $r-1$ (GRÜN [6], Satz 7). Bezeichnen wir mit $\mathfrak{G}_p^{[r-1]}$ das $r-1$-te Glied der absteigenden Zentralreihe von $\mathfrak{G}_p$, so gilt daher $\mathfrak{G}_p^{[r-1]} \leq \mathfrak{C}(\mathfrak{H})$. Es folgt also

$$\mathfrak{G}_p^{[r-1]} \leq \cap\, \mathfrak{C}(\mathfrak{H}) = \mathfrak{F}(p).$$

Nach P. HALL und G. HIGMAN [8] ist bei einer p-auflösbaren Gruppe die p-Länge durch die Klasse der p-Sylow-Gruppe beschränkt, und zwar, wenn wir mit l_p die p-Länge, mit c_p die Klasse der p-Sylow-Gruppen bezeichnen:

$c_p \geq p^{l_p - 1}$, falls $p \neq 2$ und p keine Fermat-Zahl ist ([8], S. 29, Corollar);
$c_p \geq ((p-2)^{l_p} - 1)/(p-3)$, falls $p \neq 3$ und p eine Fermat-Zahl ist;
$c_3 \geq 2^{l_3 - 1}$;
$c_2 \geq l_2$ ([8], S. 7, Theorem 1.2.6 (ii)).

In unserem Falle ergibt sich daher[7])

$$l_p(\mathfrak{G}/\mathfrak{F}(p)) \leq \begin{cases} 1 + {}_p\log(r-1) & \text{für } p \neq 2,\ p \text{ keine Fermat-Zahl} \\ {}_{p-2}\log(r-1(p-3)+1) & \text{für } p \neq 3,\ p \text{ Fermat-Zahl} \\ 1 + {}_2\log(r-1) & \text{für } p = 3 \\ r-1 & \text{für } p = 2. \end{cases}$$

Für $l_p(\mathfrak{G})$ folgen dann, wenn wir wieder r durch $r_p(\mathfrak{G})$ ersetzen, die Abschätzungen

$$l_p(\mathfrak{G}) \leq \begin{cases} 2 + {}_p\log(r_p(\mathfrak{G}) - 1) & \text{für } p \neq 2,\ p \text{ keine Fermat-Zahl} \\ 1 + {}_{p-2}\log(r_p(\mathfrak{G}) - 1)(p-3) + 1) & \text{für } p \neq 3,\ p \text{ Fermat-Zahl} \\ 2 + {}_2\log(r_3(\mathfrak{G}) - 1) & \text{für } p = 3 \\ r_2(\mathfrak{G}) & \text{für } p = 2. \end{cases}$$

Von diesen Ungleichungen sind alle für große $r_p(\mathfrak{G})$ schärfer als die Abschätzungen der Sätze 20a/b. Wir formulieren daher die Ergebnisse gesondert:

[7]) ${}_t\log r$ sei der Logarithmus von r zur Basis t.

Satz 20c. Es gilt

$$l_p(\mathfrak{G}) \leq \begin{cases} 2 + {}_p\log r_p(\mathfrak{G}) & \text{für } p \neq 2,\ p \text{ keine Fermat-Zahl} \\ 1 + {}_{p-2}\log\left(r_p(\mathfrak{G})(p-3)+1\right) & \text{für } p \neq 3,\ p \text{ Fermat-Zahl} \\ 2 + {}_2\log r_3(\mathfrak{G}) & \text{für } p = 3. \end{cases}$$

Die von P. Hall und G. Higman in [8] angegebene Abschätzung von l_p durch c_p weist zusammen mit Satz 20 auf die zentrale Bedeutung der p-Länge für die Theorie der auflösbaren Gruppen hin.

Nun behandeln wir die Verallgemeinerung der zweiten Aussage von überauflösbaren Gruppen auf auflösbare Gruppen von beliebigem Rang.

Satz 21. Sei $\mathfrak{G}$ auflösbar mit $r(\mathfrak{G}) = r$. Dann ist $\mathfrak{G}^{(k)}$ nilpotent für $k \geq \operatorname{Min}\{2r, 6\,{}_2\log r\}$.

Beweis. Ist $\mathfrak{H}$ ein Hauptfaktor der Ordnung p^n von $\mathfrak{G}$, so ist $\mathfrak{G}(\mathfrak{H})$ isomorph zu einer über $GF(p)$ irreduziblen, auflösbaren Untergruppe von $GL(n, p)$. Dabei gilt $n \leq r$. Nach Satz 9 ist $k\left(\mathfrak{G}(\mathfrak{H})\right) \leq 2r$. Beachten wir, daß $\mathfrak{G}(\mathfrak{H})$ über $GF(p)$ irreduzibel ist, daher über einer geeigneten Erweiterung von $GF(p)$ vollständig reduzibel ist in absolut irreduzible Teile, so liefert Satz 10 $k\left(\mathfrak{G}(\mathfrak{H})\right) \leq 6\,{}_2\log r$. Ist $k = \operatorname{Min}\{2r, 6\,{}_2\log r\}$, so ist also $\mathfrak{G}(\mathfrak{H})^{(k)} = \mathfrak{E}$ für alle Hauptfaktoren $\mathfrak{H}$ von $\mathfrak{G}$. Daher zentralisiert $\mathfrak{G}^{(k)}$ alle Hauptfaktoren, woraus die Behauptung $\mathfrak{G}^{(k)} \leq \mathfrak{F}$ folgt.

Einen Spezialfall betrachten wir noch gesondert:

Satz 22. Sei $\mathfrak{G}$ eine auflösbare Gruppe mit lauter abelschen Sylow-Gruppen. Haben alle Sylow-Gruppen von $\mathfrak{G}$ höchstens d Erzeugende, so gilt $\mathfrak{G}^{(k+1)} = \mathfrak{E}$ für $k \geq \operatorname{Min}\{d, 6\,{}_2\log d\}$.

Beweis. Wieder sei $\mathfrak{H}$ ein Hauptfaktor der Ordnung p^n von $\mathfrak{G}$. Nach der Voraussetzung ist $n \leq d$. Da $\mathfrak{G}_p$ abelsch ist, hat $\mathfrak{G}(\mathfrak{H})$ eine zu p teilerfremde Ordnung. Nach der Bemerkung hinter Satz 11 können wir nun Satz 11 anwenden, aber auch Satz 10. Wie im Beweis von Satz 21 ergibt sich, daß $\mathfrak{G}^{(k)} \leq \mathfrak{F}$ für $k \geq \operatorname{Min}\{d, 6\,{}_2\log d\}$. Da aber $\mathfrak{F}$ im vorliegenden Fall abelsch ist, folgt $\mathfrak{G}^{(k+1)} = \mathfrak{E}$.

Für den Spezialfall $d = 2$ hat Taunt schon früher diesen Satz bewiesen (Taunt [16]; dort die schärfere Voraussetzung, daß die Ordnung von $\mathfrak{G}$ kubusfrei ist). Die Fälle $d \leq 5$ finden sich in Huppert [9] (Satz 19 in [9]).

Wir schließen noch einen Satz über den Rang $r(\mathfrak{G})$ an.

Satz 23. Ist $\mathfrak{G}$ auflösbar, so gilt $r(\mathfrak{G}/\mathfrak{F}) \leq r_{\mathfrak{G}}\left(\mathfrak{F}/\Phi(\mathfrak{G})\right)$.

Beweis. Sei $\mathfrak{H}$ ein Hauptfaktor der Ordnung p^n von $\mathfrak{G}$ zwischen $\mathfrak{F}$ und $\Phi(\mathfrak{G})$. Dann ist $n \leq r_{\mathfrak{G}}\left(\mathfrak{F}/\Phi(\mathfrak{G})\right)$. Aus Satz 12 folgt $r\left(\mathfrak{G}(\mathfrak{H})\right) \leq n \leq r_{\mathfrak{G}}\left(\mathfrak{F}/\Phi(\mathfrak{G})\right)$; denn die dortige Voraussetzung über die vollständige Reduzibilität von $\mathfrak{G}(\mathfrak{H})$ über der algebraisch abgeschlossenen Hülle von $GF(p)$ ist nun sicher erfüllt (Hilfssatz 3). Ist $\mathfrak{T}$ der Durchschnitt der Zentralisatoren aller Hauptfaktoren von $\mathfrak{G}$ zwischen $\mathfrak{F}$ und $\Phi(\mathfrak{G})$, so bekommen wir $r(\mathfrak{G}/\mathfrak{T}) \leq r_{\mathfrak{G}}\left(\mathfrak{F}/\Phi(\mathfrak{G})\right)$. Nun haben wir noch $\mathfrak{T} = \mathfrak{F}$ nachzuweisen.

Sicher gilt $\mathfrak{D} \geq \mathfrak{F}$. Angenommen, es sei $\mathfrak{D} > \mathfrak{F}$. Dann gibt es einen Normalteiler $\mathfrak{B}$ von $\mathfrak{G}$ mit abelscher Faktorgruppe $\mathfrak{B}/\mathfrak{F}$, der in $\mathfrak{D}$ liegt. Da aber $\mathfrak{B}$ alle Hauptfaktoren von $\mathfrak{F}/\Phi(\mathfrak{G})$ zentralisiert, ist $\mathfrak{B}/\Phi(\mathfrak{G})$ nilpotent. Dann ist aber auch $\mathfrak{B}$ nilpotent (GASCHÜTZ [5], Satz 10). Daher folgt $\mathfrak{B} \leq \mathfrak{F}$, also $\mathfrak{B} = \mathfrak{F}$, entgegen unserer Annahme.

Es wäre wichtig, aus der Kenntnis der Hauptfaktoren von $\mathfrak{G}/\Phi(\mathfrak{G})$ auch auf die unterhalb von $\Phi(\mathfrak{G})$ liegenden Hauptfaktoren von $\mathfrak{G}$ zu schließen. Dies ist bisher nur unter der scharfen Bedingung $r_{\mathfrak{G}}(\mathfrak{F}/\Phi(\mathfrak{G})) = 1$ gelungen (HUPPERT [9], leichte Verallgemeinerung des dortigen Satz 13).

Literatur

[1] BRAUER, R., and C. NESBITT: On the modular representations of groups of finite order I. Univ. of Toronto Studies, math. ser. **4** (1937), 21 S. — [2] BURNSIDE, W.: On groups of order $p^a q^b$ (second paper). London Math. Soc. Proc. (2) **2**, 432—437 (1904). — [3] CLIFFORD, A. H.: Representations induced in an invariant subgroup. Annals of Math. **38**, 533—550 (1937). — [4] FRUCHT, R.: Über die Darstellung endlicher abelscher Gruppen durch Kollineationen. J. reine u. angew. Math. **166**, 16—29 (1932). — [5] GASCHÜTZ, W.: Über die Φ-Untergruppe endlicher Gruppen. Math. Z. **58**, 160—170 (1953). — [6] GRÜN, O.: Über eine Faktorgruppe freier Gruppen I. Deutsche Math. **1**, 772—782 (1936). — [7] HALL, P.: A note on soluble groups. J. London Math. Soc. **3**, 98—105 (1928). — [8] HALL, P., and G. HIGMAN: On the p-length of p-soluble groups and reduction theorems for Burnside's problem. Proc. London Math. Soc. **6**, 1—42 (1956). — [9] HUPPERT, B.: Normalteiler und maximale Untergruppen endlicher Gruppen. Math. Z. **60**, 409—434 (1954). — [10] ITÔ, N.: On the degrees of irreducible representations of a finite group. Nagoya Math. J. **3**, 5—6 (1951). — [11] ITÔ, N.: Note on A-groups. Nagoya Math. J. **4**, 79—81 (1952). — [12] JORDAN, C.: Mémoire sur les groupes résolubles. J. de Math. (7) **3**, 263—274 (1917). — [13] SCHOLZ, A., u. B. SCHOENEBERG: Einführung in die Zahlentheorie. Samml. Göschen **1955**. — [14] SCHUR, I.: Über die Darstellung der endlichen Gruppen durch gebrochene lineare Substitutionen. J. reine u. angew. Math. **127**, 20—50 (1904). — [15] SPEISER, A.: Theorie der Gruppen von endlicher Ordnung, 3. Aufl. 1937. — [16] TAUNT, D.: On A-groups. Proc. Cambridge philos. Soc. **45**, 24—42 (1949). — [17] v. d. WAERDEN, B. L.: Gruppen von linearen Transformationen. Ergebn. Math. **4/2** (1935). — [18] v. d. WAERDEN, B. L.: Moderne Algebra II, 2. Aufl. 1940. — [19] WENDT, E.: Über eine specielle Classe von Gruppen. Math. Ann. **55**, 479—492 (1902). — [20] ZASSENHAUS, H.: Über endliche Fastkörper. Abh. math. Sem. Hamburg **11**, 187—220 (1936). — [21] ZASSENHAUS, H.: Beweis eines Satzes über diskrete Gruppen. Abh. math. Sem. Hamburg **12**, 289—312 (1938). — [22] ZASSENHAUS, H.: Lehrbuch der Gruppentheorie I. 1937.

Mathematisches Institut der Universität Tübingen

(Eingegangen am 18. Februar 1957)